

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q	A2	(11) International Publication Number: WO 97/18323 (43) International Publication Date: 22 May 1997 (22.05.97)
(21) International Application Number: PCT/US96/19337 (22) International Filing Date: 18 November 1996 (18.11.96) (30) Priority Data: 08/558,666 16 November 1995 (16.11.95) US 08/007,312 27 February 1996 (27.02.96) US 08/725,304 21 October 1996 (21.10.96) US (71) Applicant (for all designated States except US): ICOS CORPORATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA 98021 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): HOEKSTRA, Merl, F. [US/US]; 10321 216th Street, S.E., Snohomish, WA 98290 (US). HOLTZMAN, Doug, A. [US/US]; 4308 4th Avenue, N.E., Seattle, WA 98105 (US). KEEGAN, Kathleen, S. [US/US]; 5812 Mercer Way, Mercer Island, WA 98040 (US). (74) Agent: NOLAND, Greta, E.; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606-6402 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: CELL CYCLE CHECKPOINT PIK-RELATED KINASE MATERIALS AND METHODS (57) Abstract The present invention generally relates to genes encoding cell cycle checkpoint phosphatidylinositol kinase (PIK)-related proteins essential to DNA damage responses in cells. These PIK-related kinases are required in regulatory pathways that arrest the cell cycle following DNA damage to allow DNA repair prior to mitosis or initiation of DNA replication. More particularly, the invention provides a novel human cell cycle checkpoint PIK-related kinase, MCCS1, and polynucleotide sequences encoding the MCCS1. Assays for identifying modulators of MCCS1 useful as, for example, chemotherapy and radiation adjuvants, are also provided by the invention. Further, assays for identifying modulators of the cell cycle checkpoint phosphatidylinositol kinase (PIK)-related protein identified as ATM are provided.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

CELL CYCLE CHECKPOINT PIK-RELATED KINASE MATERIALS AND METHODS

FIELD OF THE INVENTION

5 The present invention generally relates to genes encoding cell-cycle
checkpoint phosphatidylinositol kinase (PIK)-related genes and proteins essential to
DNA damage responses in cells. The checkpoint kinases play a role in the
surveillance of DNA damage that occurs as a result of replication errors, DNA
mismatches, radiation treatment, or chemotherapeutic drugs. These kinases are
required in regulatory pathways that lead to cell cycle arrest following DNA damage,
10 giving the cell notice and time to correct lesions prior to the initiation of DNA
replication. More particularly, the invention relates to a novel human PIK-related
kinase, Mammalian Cell Cycle Surveillance 1 (MCCS1), polynucleotides encoding
the PIK-related kinase, and methods for assaying and modulating the enzymatic
activity of the kinase and related kinases.

15

BACKGROUND

The process of eukaryotic cell growth and division is the somatic or
mitotic cell cycle which consists of four phases, the G₁ phase, the S phase, the G₂
phase, and the M phase. The G₁, S, and G₂ phases are collectively referred to as
interphase of the cell cycle. The cell cycle is structurally and functionally conserved
20 in its basic process and mode of regulation across all eukaryotic species. During the
G₁ (gap) phase, biosynthetic activities of the cell progress at a high rate. The S
(synthesis) phase begins when DNA synthesis starts and ends when the DNA content
of the nucleus of the cell has been replicated and two identical sets of chromosomes
are formed. The cell then enters the G₂ (gap) phase which continues until mitosis
25 starts. In mitosis, the chromosomes pair and separate and two new nuclei form, and
in cytokinesis the cell itself splits into two daughter cells each receiving one nucleus

-2-

containing one of the two sets of chromosomes. Mitosis and cytokinesis together form the M (mitosis) phase of the cell cycle. Cytokinesis terminates the M phase and marks the beginning of interphase of the next cell cycle. The sequence in which the events in the cell cycle proceed is tightly regulated such that the initiation of one cell cycle event is dependent on the completion of the prior cell cycle event. This allows fidelity in the duplication and segregation of genetic material from one generation of cells to the next.

The term "cell cycle checkpoints" refers to the proteins, signals, processes, and feedback controls that integrate discontinuous events during cellular replication, in order to maintain essential dependencies within the cell cycle. The present invention specifically relates to the cell cycle checkpoint that ensures that mitosis is delayed until the completion of DNA synthesis and/or the accurate repair of DNA damage occurs.

Failure of cell cycle checkpoints predisposes individuals to or directly causes many disease states such as cancer, ataxia telangiectasia, embryo abnormalities, and various immunological defects associated with aberrant B and T cell development. The latter are associated with pathological states such as lupus, arthritis and autoimmune diseases. Intense research efforts have therefore focused on identifying cell cycle checkpoints and the proteins essential for the function of the checkpoints.

Genetic analysis in the yeasts *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* has identified a number of genes important for cell cycle arrest and DNA repair responses to ionizing radiation (IR). For a review, see Carr and Hoekstra, *Trends in Cell Biology*, 5: 32-40 (1995). One such gene, identified in both yeasts, is required for a DNA damage checkpoint which arrests the cell cycle at the G2 phase, as well as a related checkpoint which monitors the completion of DNA synthesis and arrests the cell cycle at the S phase. The gene is named *rad3+* in *S. pombe* [Seaton *et al.*, *Gene*, 119: 83-89 (1992)], *MEC1/ESR1* in *S. cerevisiae* [Kato *et al.*, *Nuc. Acids. Res.*, 22(15): 3104-3112 (1994)], and is hereinafter referred to as *rad3+*. Cells having mutations in *rad3+* fail to either sense or appropriately respond to DNA damage and subsequently lose viability more rapidly than wild type cells after exposure to clastogenic agents or events (*e.g.*, IR, DNA damaging agents,

-3-

and mutations affecting chromosomal integrity). See Weinert *et al.*, *GENES & DEVELOPMENT*, 8: 652-665 (1994) and Al-Khodairy *et al.*, *EMBO J.*, 11(4): 1343-1350 (1992). This sensitivity to IR (radiosensitivity) can be caused by defects in checkpoint responses or defects in direct DNA repair reactions.

5 The product of the *rad3+* gene is an approximately 270 kD protein that falls into a growing family of high molecular weight PIK-related kinases. See Hunter, *Cell*, 83: 1-4 (1995) for a discussion of this family of kinases. The primary structures of the catalytic domains found in members of this gene family are closely related to well characterized phosphatidylinositol kinases. This structural relationship
10 initially suggested that these PIK-related kinases might be capable of phosphorylating lipids. When the substrate specificity of the PIK-related kinases is examined, however, these enzymes appear to function as protein kinases and have yet to be demonstrated to phosphorylate phosphatidylinositides. Hartley *et al.*, *Cell*, 82: 849-856 (1995) reports that purified preparations highly active in protein kinase assays
15 failed to show lipid kinase activity. Additional PIK-related kinases identified include: the TEL1 gene product from *S. cerevisiae* which affects telomere length [Greenwell *et al.*, *Cell*, 82: 823-829 (1995)], and *Mei41+* gene product from *Drosophila melanogaster* which is important for a G2 checkpoint and meiotic development [Hari *et al.*, *Cell*, 82: 815-821 (1995)], the DNA-PK gene product from mouse which is
20 important in immunoglobulin rearrangements and processing of DNA double strand breaks, and the FRAP gene product which is important in the G1/S transition [Brown, E. *et al.*, *Nature*, 377:441-446 (1995)]. Mutations in the DNA-PK gene can result in the Severe Combined Immunodeficiency Syndrome (SCID) defect (Hartley *et al.*, *supra*).

25 In humans, less is known about the molecular components required for checkpoint function. One component of the mammalian checkpoint machinery has been identified through the analysis of the human disease syndrome ataxia-telangiectasia (AT). Patients with AT show a diverse set of clinical symptoms, including predisposition to a variety of tumor types. Fibroblasts from AT patients are
30 radiosensitive and fail to undergo cell cycle arrest following treatment with IR leading to a phenomenon termed radioresistant DNA synthesis. This is reminiscent of the *S. pombe rad3* defect where cells fail to sense or respond appropriately to DNA damage.

-4-

Interestingly, the locus responsible for AT, the Ataxia-Telangiectasia Mutated (ATM) gene, was recently described in Savitsky *et al.*, *Science*, 268: 1749-1753 (1995) and the partial cDNA encodes a protein with amino acid similarity to the *rad3+* gene. Savitsky *et al.*, *Human Molecular Genetics*, 4(11):2025-2032 (1995) describes isolation of a cDNA encoding full length ATM. The increased radiosensitivity of *rad3+* yeast mutants and of mammalian cells lacking functional ATM suggests that these proteins may comprise a family of checkpoint proteins.

Kuerbitz *et al.*, *Proc. Natl. Acad. Sci. USA*, 89: 7492-7495 (1992) establishes that the tumor suppressor p53 is required for a G1 checkpoint and cell cycle arrest observed following DNA damage. Irradiation of cells results in increased levels of p53 leading to the transcriptional activation of p53 responsive genes. One such p53-induced target is the product of the WAF1 gene (also called p21, CIP1, and sdil). WAF1 is a member of an expanding class of cell cycle regulators termed cyclin-dependent kinase inhibitory proteins. The activities of cyclin-dependent kinases control transit through the cell cycle. Transcriptional activation of WAF1 thus provides a direct link between DNA damage-dependent induction of p53 and the inhibition of kinases essential for cell cycle progression. See Elledge and Harper, *Current Opinion in Cell Biology*, 6: 847-852 (1994). Mutations in the p53 gene are one of the most common genetic alterations in human cancers. For example, Baker *et al.*, *Science*, 244:217-221 (1989) reports that approximately 70% of human colorectal carcinomas contain deletions or mutant copies of the p53 gene. In addition, Fearon *et al.*, *Cell*, 61: 759-767 (1990) reports that breast, lung, bladder and brain tumors have been associated with loss of chromosome 17p, the region to which the p53 gene localizes.

At present there is relatively little known about the molecular components of the G2 checkpoints in mammalian cells. Caffeine is a chemical entity which abrogates G2 checkpoint control. Russell *et al.*, *Cancer Res.*, 55: 1639-1642 (1995) and Powell *et al.*, *Cancer Res.*, 55: 1643-1648 (1995) report that analysis of cell lines which differ only by the presence or absence of functional p53 demonstrated preferential caffeine-enhanced sensitization to IR in those cells lacking the p53-dependent G1 checkpoint. Thus, the conversion of potentially lethal damage into

-5-

lethal damage is greater in cells lacking the G1 and G2 checkpoints in comparison to cells containing an intact G1 checkpoint.

While certain cells undergo DNA damage-dependent cell cycle arrest, other cells appear to respond to DNA damage by initiating an intrinsic suicide program termed apoptosis or programmed cell death. The factors determining which process occurs are not fully understood. Recent work has demonstrated an important role for p53 both in the regulation of G1 cell cycle transitions and apoptosis. Symonds *et al.*, *Cell*, 78: 703-711 (1994) describe p53-dependent apoptosis as suppressing tumor growth and progression *in vivo*.

High doses of radiation and chemotherapy are used to treat tumor cells in order to damage DNA so severely that the cells will die. However, even though tumor cells having mutations in the p53 gene are defective in a G1 checkpoint, they can still repair DNA damaged induced by radiation or chemotherapy. The present invention contemplates, for example, that inhibition of the G2 checkpoint in tumor cells should lead to a state in which tumor cells are incapable of repairing DNA damage therefore sensitizing the tumor cells to DNA damaging agents. Normal cells, containing intact G1 and G2 checkpoints, should still be able to repair DNA damage in the presence of a G2 checkpoint-specific inhibitor. Thus, treatment of tumors with a G2 checkpoint-specific inhibitor followed by radiation or chemotherapy should increase the efficacy of cell killing and thereby decrease the required doses of toxic DNA-damaging agents.

There thus exists a need in the art for identification of the mammalian proteins that are involved in the cell cycle checkpoints in order to develop therapies for the human disease states associated with defective cell cycle checkpoints and for the isolation of the genes encoding those proteins which in themselves may be useful as therapeutics or which would enable the development of therapeutically useful modulators of the proteins encoded by the genes.

SUMMARY OF THE INVENTION

The present invention provides novel human PIK-related kinases essential for a cell cycle checkpoint that responds in the G2 phase of the cell cycle to both damaged and unreplicated DNA.

-6-

In one of its aspects, the present invention provides purified and isolated polynucleotides (*e.g.*, DNAs and RNAs, both coding and non-coding strands thereof) encoding the cell cycle checkpoint PIK-related kinase M CCS1 and polynucleotides encoding other cell cycle checkpoint PIK-related kinases that exhibit
5 about 50, about 60, or about 65 % nucleotide identity to the M CCS1 polynucleotide region encoding the M CCS1 kinase domain (M CCS1 α nucleotides 6579 to 7562 of SEQ ID NO: 30 or M CCS1 β nucleotides 6457 to 7440 of SEQ ID NO: 32). Alternatively, the M CCS1-like PIK-related kinases exhibit about 40%, about 45 %, or about 50% amino acid identity to the M CCS1 kinase domain (M CCS1 α amino
10 acids 2083 to 2410 of SEQ ID NO: 31 or M CCS1 β amino acids 2152 to 2480 of SEQ ID NO: 33). Polynucleotides contemplated by the invention include genomic DNAs, RNAs, cDNAs and wholly or partially chemically synthesized DNAs. Preferred polynucleotides of the invention comprise the M CCS1 α DNA sequence set out in SEQ ID NO: 30, the partial M CCS1 β DNA sequence set out in SEQ ID NO: 3, the
15 full length M CCS1 β DNA sequence set out in SEQ ID NO: 32, and DNA sequences which hybridize to the noncoding strands thereof under stringent conditions or which would hybridize but for the redundancy of the genetic code. Exemplary stringent hybridization conditions are as follows: hybridization at 65 °C in 3X SSC, 20mM NaPO₄, pH 6.8 and washing at 65 °C in 0.2X SSC. It is understood by those of skill
20 in the art that variation in these conditions occurs based on the length and GC nucleotide base content of the sequences to be hybridized. Formulas standard in the art are appropriate for determining exact hybridization conditions. See Sambrook *et al.*, 9.47-9.51 in *Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). The M CCS1 α DNA of SEQ ID NO: 30 was
25 deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852, on November 3, 1995 as an insert in plasmid pBSHFB/HT2-27 in *E. coli* DH5 α and was assigned ATCC Accession No. 69951. The M CCS1 β DNA of SEQ ID NO: 32, was deposited with the ATCC on November 7, 1995 as an insert in plasmid 517 in *E. coli* DH5 α and was assigned ATCC
30 Accession No. 69950.

The DNA sequence information provided by the present invention makes possible the identification and isolation of DNAs encoding related molecules

-7-

by well-known techniques such as DNA/DNA hybridization as described above and polymerase chain reaction (PCR) cloning. As one series of examples, knowledge of the sequence of a cDNA encoding MCCS1 makes possible the isolation by DNA/DNA hybridization of genomic DNA sequences encoding the kinase and expression control regulatory sequences such as promoters, operators and the like. Similarly, knowledge of a partial cDNA sequence encoding MCCS1 β make isolation of a complete cDNA possible. DNA/DNA hybridization procedures carried out with DNA sequences of the invention under stringent conditions are likewise expected to allow the isolation of DNAs encoding allelic variants of the PIK-related kinase; non-human species enzymes homologous to the PIK-related kinase; and other structurally related proteins sharing one or more of the enzymatic activities, or abilities to interact with members or regulators, of the cell cycle checkpoint pathway in which MCCS1 participates. Polynucleotides of the invention when detectably labelled are also useful in hybridization assays to detect the capacity of cells to synthesize MCCS1. The DNA sequence information provided by the present invention also makes possible the development, by homologous recombination or "knockout" strategies [see, Capecchi, *Science*, 244: 1288-1292 (1989)], of rodents that fail to express functional MCCS1 or that express a variant of MCCS1. Such rodents are useful as models for studying the activities of MCCS1 and MCCS1 modulators *in vivo*. Polynucleotides of the invention may also be the basis for diagnostic methods useful for identifying a genetic alteration(s) in the MCCS1 locus that underlies a disease state or states. Also made available by the invention are anti-sense polynucleotides relevant to regulating expression of MCCS1 by those cells which ordinarily express the same.

The invention also provides autonomously replicating recombinant constructions such as plasmid and viral DNA vectors incorporating polynucleotides of the invention, especially vectors in which the polynucleotides are functionally linked to an endogenous or heterologous expression control DNA sequence and a transcription terminator.

According to another aspect of the invention, host cells, especially unicellular host cells such as procaryotic and eukaryotic cells, are stably transformed or transfected with DNAs of the invention in a manner allowing expression of the PIK-related kinase therein. Host cells of the invention are conspicuously useful in

methods for the large scale production of MCCS1 wherein the cells are grown in a suitable culture medium and the desired enzymes are isolated from the cells or from the medium in which the cells are grown.

5 MCCS1 products having part or all of the amino acid sequence set out in SEQ ID NO: 31, SEQ ID NO: 4, or SEQ ID NO: 33 are contemplated. Use of mammalian host cells is expected to provide for such post-translational modifications (e.g., myristoylation, glycosylation, truncation, lipidation and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. The enzyme products of the invention may be full length polypeptides, fragments or variants. Variants comprise 10 MCCS1 products wherein one or more of the specified (*i.e.*, naturally encoded) amino acids is deleted or replaced or wherein one or more nonspecified amino acids are added: (1) without loss of the kinase activity specific to MCCS1; or (2) with disablement of the kinase activity specific to MCCS1; or (3) with disablement of the ability to interact with members or regulators of the cell cycle checkpoint pathway. 15 Substrates of MCCS1 and proteins which interact with MCCS1 may be identified by various assays.

Substrates of MCCS1 may be identified by incorporating test compounds in assays for kinase activity. MCCS1 kinase is resuspended in kinase 20 buffer and incubated either in the presence or absence of the test compound (e.g., casein, histone H1, or appropriate substrate peptide). Moles of phosphate transferred by the kinase to the test compound are measured by autoradiography or scintillation counting. Transfer of phosphate to the test compound is indicative that the test compound is a substrate of the kinase.

25 Interacting proteins may be identified by the following assays.

A first assay contemplated by the invention is a two-hybrid screen. The two-hybrid system was developed in yeast [Chien *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582 (1991)] and is based on functional *in vivo* reconstitution of a transcription factor which activates a reporter gene. Specifically, a polynucleotide 30 encoding a protein that interacts with MCCS1 is isolated by: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA

binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding a first fusion of part or all of MCCS1 and either the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a library of second hybrid DNA sequences encoding second fusions of part or all of putative MCCS1 binding proteins and the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion; detecting binding of an MCCS1 interacting protein to MCCS1 in a particular host cell by detecting the production of reporter gene product in the host cell; and isolating second hybrid DNA sequences encoding the interacting protein from the particular host cell. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

Other assays for identifying proteins that interact with MCCS1 may involve immobilizing MCCS1 or a test protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the amount of label bound. Bound label indicates that the test protein interacts with MCCS1.

Another type of assay for identifying MCCS1 interacting proteins involves immobilizing MCCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling a test protein with a compound capable of exciting the fluorescent agent, contacting the immobilized MCCS1 with the labelled test protein, detecting light emission by the fluorescent agent, and identifying interacting proteins as test proteins which result in the emission of light by the fluorescent agent. Alternatively, the putative interacting protein may be immobilized and MCCS1 may be labelled in the assay.

Also comprehended by the present invention are antibody products (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, CDR-grafted antibodies and the like) and other binding proteins (such as those identified in the assays above) which are specific for the MCCS1 kinases of the invention. Binding proteins can be developed using isolated natural or recombinant enzymes. The binding proteins are useful, in turn, for purifying recombinant and

-10-

naturally occurring enzymes and identifying cells producing such enzymes. Specifically illustrating monoclonal antibodies of the invention are the monoclonal antibodies produced by hybridoma cell lines 224C and 224F which were deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD 20852 on November 7, 1996 and assigned ATCC Accession Nos. HB 12233 and HB 12234, respectively. Assays for the detection and quantification of proteins in cells and in fluids may involve a single antibody substance or multiple antibody substances in a "sandwich" assay format. The binding proteins are also manifestly useful in modulating (*i.e.*, blocking, inhibiting, or stimulating) enzyme/substrate or enzyme/regulator interactions. Anti-idiotypic antibodies specific for PIK-related kinase binding proteins are also contemplated.

The invention contemplates that mutations in the M CCS1 gene that result in loss of normal function of the M CCS1 gene product underlie human disease states in which failure of the G₂ cell cycle checkpoint is involved. Gene therapy to restore M CCS1 activity would thus be indicated in treating those disease states (for example, testicular cancer). Delivery of a functional M CCS1 gene to appropriate cells is effected *in vivo* or *ex vivo* by use of viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus) or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). For reviews of gene therapy technology see Friedmann, *Science*, 244: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature*, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states preventing the expression of or inhibiting the activity of M CCS1 will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of M CCS1. Antisense nucleic acids (preferably 10 to 20 base pair oligonucleotides) capable of specifically binding to M CCS1 expression control sequences or M CCS1 RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the M CCS1 target sequence in the cell and prevents transcription or translation of the target sequence. Phosphothioate and methylphosphate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense

-11-

oligonucleotides may be further modified by poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end.

Moreover, for example, if a particular form of cancer results from a mutation in a gene other than MCCC1 such as the p53 gene, an agent which inhibits the transcription or the enzymatic activity of MCCC1 and thus the G₂ cell cycle checkpoint may be used to render cancerous cells more sensitive to chemotherapy or radiation therapy. The therapeutic value of such an agent lies in the fact that current radiation therapy or chemotherapy in most cases does nothing to overcome the ability of the p53 mutant cancerous cell to sense and correct the DNA damage imposed as a result of the treatment. As a result, a cancer cell can simply repair the DNA damage. Modulating agents of the invention may therefore be chemotherapy and radiation adjuvants or may be directly active as chemotherapeutic drugs themselves.

Agents that modulate MCCC1 kinase activity may be identified by incubating a test compound with MCCC1 immunopurified from cells naturally expressing the PIK-related kinase, with MCCC1 obtained from recombinant procaryotic or eukaryotic host cells expressing the enzyme, or with purified MCCC1, and then determining the effect of the test compound on MCCC1 activity. The activity of the PIK-related kinase can be measured by determining the moles of ³²P-phosphate transferred by the kinase from gamma-³²P-ATP to either itself (autophosphorylation) or to an exogenous substrate such as a lipid or protein. The amount of phosphate incorporated into the substrate is measured by scintillation counting or autoradiography. An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCC1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCC1 kinase. In another aspect, agents that modulate both MCCC1 and ATM or modulate one of the enzymes are also contemplated. Agents which modulate MCCC1 are screened in a kinase assay as described above in which ATM is the phosphorylating enzyme.

-12-

In a presently preferred assay, a MCCC1-specific antibody linked to agarose beads is incubated with a cell lysate prepared from host cells expressing the kinase. The beads are washed to remove proteins binding nonspecifically to the beads and the beads are then resuspended in kinase buffer. The reaction is initiated by the addition of gamma-³²P-ATP and an appropriate exogenous substrate such as lipid or peptide. The activity of the kinase is measured by determining the moles of ³²P-phosphate transferred either to the kinase itself or the added substrate. In a preferred embodiment the host cells lack endogenous MCCC1 and/or ATM kinase activity. The selectivity of a compound that modulates the kinase activity of MCCC1 can be evaluated by comparing its activity on MCCC1 to its activity on other known PIK-related kinases. The combination of the recombinant MCCC1 products of the invention with other recombinant PIK-related kinase products in a series of independent assays provides a system for developing selective modulators of MCCC1.

Furthermore, combinatorial libraries, peptide and peptide mimetics, defined chemical entities, oligonucleotides, and natural product libraries may be screened for activity as modulators in assays such as those described below.

For example, an assay for identifying modulators of MCCC1 kinase activity involves incubating an MCCC1 kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate, both in the presence and absence of a test compound, and measuring the moles of phosphate transferred to the substrate. An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCC1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCC1 kinase.

Moreover, assays for identifying compounds that modulate interaction of MCCC1 with other proteins may involve: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA-binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding

-13-

a first fusion of part or all of M CCS1 and the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a second hybrid DNA sequence encoding part or all of a protein that interacts with M CCS1 and the DNA binding domain or activating domain of the transcription factor which is not
5 incorporated in the first fusion; evaluating the effect of a test compound on the interaction between M CCS1 and the interacting protein by detecting binding of the interacting protein to M CCS1 in a particular host cell by measuring the production of reporter gene product in the host cell in the presence or absence of the test compound; and identifying modulating compounds as those test compounds altering
10 production of the reported gene product in comparison to production of the reporter gene product in the absence of the modulating compound. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

15 Another type of assay for identifying compounds that modulate the interaction between M CCS1 and an interacting protein involves immobilizing M CCS1 or a natural M CCS1 interacting protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the effect of a test compound on the amount of label bound wherein a reduction in the label
20 bound in the present of the test compound compared to the amount of label bound in the absence of the test compound indicates that the test agent is an inhibitor of M CCS1 interaction with protein. Conversely, an increase in the bound in the presence of the test compound compared to the amount label bound in the absence of the compound indicates that the putative modulator is an activator of M CCS1
25 interaction with the protein.

Yet another method contemplated by the invention for identifying compounds that modulate the binding between M CCS1 and an interacting protein involves immobilizing M CCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling the interacting protein with a
30 compound capable of exciting the fluorescent agent, contacting the immobilized M CCS1 with the labelled interacting protein in the presence and absence of a test compound, detecting light emission by the fluorescent agent, and identifying

-14-

modulating compounds as those test compounds that affect the emission of light by the fluorescent agent in comparison to the emission of light by the fluorescent agent in the absence of the test compound. Alternatively, the M CCS1 interacting protein may be immobilized and M CCS1 may be labelled in the assay.

5 The present invention further provides a cell-based complementation assay for identifying compounds which modulate the activity of M CCS1 or ATM. The assay involves complementation of a phenotypic trait associated with a genetic alteration in the cell. For example, the genetic alteration identified as *esr1-1* results in cellular sensitivity to DNA damage in yeast cells [Kato *et al.*, *Nuc. Acids. Res.*,
10 22(15): 3104-3112 (1994)]. *esr1-1* cells fail to either sense or appropriately response to DNA damage after exposure to DNA damaging agents such as ionizing radiation or clastogenic agents. The phenotypic trait of the genetically altered cell is complemented by transforming and expressing M CCS1 or ATM in the cell. The transformed cells are exposed to DNA damaging treatment (e.g. ionizing radiation)
15 in the presence and absence of a test compound and sensitivity of the cells to DNA damage is measured. Agents that affect the cell sensitivity to DNA damaging activity of M CCS1 and/or ATM are identified as modulators.

 Modulators of M CCS1 may affect its kinase activity, its localization in the cell, and/or its interaction with members of the cell cycle checkpoint pathway.
20 M CCS1 modulators may be formulated in compositions comprising pharmaceutically acceptable carriers. Such compositions may additionally include chemotherapeutic agents. Dosage amounts indicated would be sufficient to result in modulation of M CCS1 activity *in vivo*. Selective modulators may include, for example, polypeptides or peptides which specifically bind to M CCS1 or M CCS1 nucleic acid,
25 oligonucleotides which specifically bind to the PIK-related kinase or PIK-related kinase nucleic acid, and/or other non-peptide compounds (e.g., isolated or synthetic organic molecules) which specifically react with M CCS1 or M CCS1 nucleic acid. Mutant forms of M CCS1 which affect the enzymatic activity or cellular localization of wild-type M CCS1 are also contemplated by the invention. Presently preferred
30 regions of the PIK-related kinases which are targets for the development of selective modulators include, for example, the following four regions: the M CCS1 α amino terminal effector domain (amino acids 1 to 1081 of SEQ ID NO: 31), the M CCS1 β

-15-

amino terminal effector domain (amino acids 1 to 1150 of SEQ ID NO: 33), the
MCCS1 α *rad3*+ domain (amino acids 1082 to 2082 of SEQ ID NO: 31), the
MCCS1 β *rad3*+ domain (amino acids 1151 to 2151 of SEQ ID NO: 33), the
MCCS1 α PIK domain (amino acids 2083 to 2410 of SEQ ID NO: 31), and the
5 MCCS1 β PIK domain (amino acids 2152 to 2480 of SEQ ID NO: 33).

-16-

DETAILED DESCRIPTION

The present invention is illustrated by the following examples. Example 1 details the isolation of cDNAs encoding M CCS1 kinases. Example 2 describes mapping of the human M CCS1 gene to human chromosome 3. The recombinant expression of M CCS1 in *E. coli* and insect cells is respectively described in Examples 3 and 4. Example 4 also presents assays for measuring M CCS1 kinase activity. Example 5 describes the production of M CCS1-specific polyclonal and monoclonal antibodies. Example 6 reports the immunoprecipitation of M CCS1 kinase associated activity from mouse testes. Example 7 examines the expression of M CCS1 mRNA in various human tissues and cancer cell lines. Example 8 describes analyses of M CCS1 mRNA and protein expression in mouse testes. Example 9 describes analyses of M CCS1 protein expression in meiotic cells. Assays for substrates and interacting proteins of M CCS1 are described in Example 10. Example 11 describes modulators and assays for modulators of the kinase activity of M CCS1. Example 12 describes the cell-based complementation assay for identifying modulators of M CCS1 and/or ATM and Example 13 describes the kinase activity of ATM.

Example 1

cDNAs encoding the PIK-related kinase M CCS1 were isolated by a series of PCR reactions.

An alignment of the amino acid sequences of *S. pombe rad3+* (Hari *et al.*, *supra*) and *S. cerevisiae MEC1* (Kato *et al.*, *supra*) was the basis for design of seven degenerate oligonucleotides that encoded (or were complementary to) the regions of highest homology/lowest degeneracy between the sequences and contained convenient restriction sites to facilitate cloning of amplification products. The oligonucleotides were then used in a PCR-based assay to isolate a related human sequence.

Initially, PCR amplifications were performed on cDNA preparations from rat T-cells, human peripheral blood mononuclear cells (PBMC), and *S. cerevisiae* genomic DNA. Five oligonucleotide pairs were used (oDH15a/oDH16, oDH15b/oDH16, oDH17a/oDH16, oDH15a/oDH17b, and oDH15b/oDH17b) for the primary amplifications. The sequences of the oligonucleotide primers included

-17-

inosines and are set out below in IUPAC nomenclature for degenerate nucleotide positions.

oDH15a (SEQ ID NO: 5)

5' GCA GAC GGA TCC GGI WCI GAY GGI AAY HTI TAY 3'

5

oDH15b (SEQ ID NO: 6)

5' GCA GAC GGA TCC GGI WCI GAY GGI AAY 3'

oDH16 (SEQ ID NO: 7)

5' GCA GAC GAA TTC RCA RTY RAA RTC IAC RTG 3'

oDH17a (SEQ ID NO: 8)

10

5' GCA GAC GGA TCC AAR TTY

CCI CCI RTI YTI TAY SAR TGG TT 3'

oDH17b (SEQ ID NO: 9)

5' GCA GAC GAA TCC AAC CAY

TSR TAI ARI AYI GGI GGR AAY TT 3'

15 PCR was performed on reaction mixtures of 1X PCR buffer (Perkin Elmer Cetus, Emeryville, California), 2-3 μ M oDH primers, 1.5mM MgCl₂, 200 μ M dNTPs, and 0.5 μ l Amplitaq polymerase. The reaction was performed in a Perkin-Elmer Cetus Thermocycler Model 480 under the following conditions: denaturation at 94°C for 1 minute, annealing at 64°C for 2 minutes, and elongation at 72°C for 1 minute for 3 cycles. The procedure was then repeated using 60°C annealing temperature for 3 cycles, 56°C annealing for 3 cycles, and finished with denaturation at 94°C for 1 minute, annealing at 54°C (2 minutes, and elongation at 72°C for 1 minute for 30 cycles. PCR products were separated on 2 or 4% Tris Acetate EDTA (TAE) agarose gels, stained with ethidium bromide, and DNA products were visualized by UV
20
25 fluorescence.

-18-

From the primary amplifications of yeast genomic DNA, rat T-cell cDNA, and human PBMC cDNA, only a single reaction with yeast genomic DNA (oDH17a/oDH16) gave a visible amplification product, resulting in a product that was the expected size for the region of the *S. cerevisiae* *MEC1* gene between these primers. Further analysis of the oDH17a/oDH16 amplifications that utilized rat T-cell and PBMC cDNA was therefore performed. To remove oligonucleotides and "primer dimers" that might interfere with subsequent PCR, primary reactions were purified prior to reamplification.

A "nested" PCR strategy was employed, and amplifications were repeated with primer pairs oDH18a/oDH16 and oDH18b/oDH16 under reaction conditions described above with cycle times of denaturation of 94°C for 1 minute, annealing at 55°C for 1 minute, and elongation at 72°C for 30 seconds for 30 cycles. The sequences of the oDH18a and oDH18b oligonucleotide primers included inosines and are set out below in IUPAC nomenclature for degenerate nucleotide positions.

oDH18a (SEQ ID NO: 10)

5' GCA GAC GGA TCC YTI GGI YTI GGI GAY CGI CA 3'

oDH18b (SEQ ID NO: 11)

5' GCA GAC GGA TCC YTI GGI YTI GGI GAY AGR CA 3'

An approximately 90 base pair (bp) product (the expected size amplification product for these primers) was seen in the reamplifications of the yeast genomic and human PBMC cDNA primary reactions. No 90 bp product was seen in the reamplification of the primary reaction on rat T-cell cDNA and this reaction was not analyzed further.

In addition to the approximately 90 bp product, several other non-specific bands were also present, though significantly fewer than were observed when the primary reactions were reamplified with oDH17a/oDH16. While the approximately 90 bp product was present in both the oDH18a/oDH16 and oDH18b/oDH16 reamplifications of the yeast genomic DNA primary reactions, only the oDH18a/oDH16 reaction yielded the appropriate size fragment during

-19-

reamplification of the human PMBC cDNA primary reaction. This was presumed to reflect codon usage in the human gene (compare primers oDH18a and oDH18b). The approximately 90 bp product from the oDH18a/oDH16 reamplification of the human PMBC cDNA primary reaction was gel purified and subcloned into the pBluescript SKII+ cloning vector (Stratagene, La Jolla, California) and sequenced.

Analysis of the sequence encoded by the 90 bp product indicated that the deduced amino acid sequence was similar to both *S. cerevisiae* MEC1 and *S. pombe rad3+*, but was not identical to either. To identify a larger region of coding sequence and extend the sequence comparison, a non-degenerate oligonucleotide, oDH23 5' GACGCAGAATTCAACCAGTCAAAGAATCAAAGAG 3' (SEQ ID NO: 12), was synthesized for use in additional amplification reactions. Reamplification of the purified PMBC cDNA primary reaction with oDH17a/oDH23 led to the production of an amplification product of 174 bp. This fragment was then purified, subcloned and sequenced as described above. Computer analysis of the conceptual translation product confirmed its relationship (similar but not identical) to MEC1 and *rad3+*. This PCR fragment was then used as a probe to screen a plasmid library containing macrophage cDNA using the following hybridization conditions: incubation of nitrocellulose filters with radiolabelled probes in 3X SSC, 5X Denhardt's, 0.1% sarcosyl, 20mM NaPO₄ pH 6.8, 100 ug/ml single stranded salmon sperm DNA, for 18 to 24 hours at 65°C. Washes were done 3 times in 0.2X SSC, 0.1% SDS at 65°C for 30 minutes (with changes of wash buffer). Four positive clones were isolated, and the nucleotide sequence of each was determined. Computer analysis of the four sequences demonstrated that they were overlapping clones derived from a locus with homology to the *rad3+* gene from *S. pombe*. Clone 517 (ATCC 69950) contained a 2.8 kbp insert and its DNA and deduced amino acid sequence are set out in SEQ ID NOs: 3 and 4, respectively. The clone contained an open reading frame encoding an amino terminal truncated protein product of 870 amino acids which were 39% identical to the COOH-terminus of *rad3+*. The protein product of the cDNA insert was named MCCS1β.

The sequence of clone 517 was used to design the oligonucleotides, mo3 5'-CTACAGAGCCAAGGAG-3' (SEQ ID NO: 13) and mo6 5'-TCGAGCTATGCTACTAGTGGGC-3' (SEQ ID NO: 14), which were used to

-20-

generate a probe using a gel purified EcoRI fragment derived from clone 517 as a template. The PCR conditions were as follows: 50 ng DNA fragment, 1X PCR buffer (Perkin-Elmer Cetus), 1.5mM MgCl₂, 200μM dATP, dGTP, and TTP, 1μM dCTP, 50μCi α³²P-dCTP, 10ng/ml each oligonucleotide, 1U AmpliTaq (Perkin-Elmer Cetus). The reaction was performed in a Perkin-Elmer Cetus Thermocycler Model 480 for an initial denaturing cycle at 94°C for 4 minutes followed by 20 cycles of 94°C for 15 seconds, 60°C for 15 seconds, 72°C for 30 seconds. Unincorporated nucleotides were removed using a Stratagene Nuc-trap Push Column.

Since Northern blot analyses showed that the expression of the mRNA corresponding to clone 517 was highest in testis, one million clones from a human testis cDNA library (Stratagene #939202) were screened with the PCR-generated probe and eleven clones were obtained. The two longest clones, HT2 and HT9, were chosen for analysis. HT2 contained a 4.7 Kb insert (corresponding to nucleotide 2974 of SEQ ID NO: 30 and extending further downstream than SEQ ID NO: 1) and HT9 contained a 5485 bp insert (corresponding to nucleotides 2152 to 7624 of SEQ ID NO: 30). Nucleotide sequence analysis revealed that in the region common to both cDNA clones there was a single base pair insertion of a T at nucleotide 3233 in HT9. This nucleotide insertion causes the predicted amino acid reading frame to shift and then terminate and is believed to be an error introduced by reverse transcriptase in clone HT9.

In order to isolate a clone containing an additional 2.5 Kb, one million clones from each of three additional cDNA libraries were screened: a human fetal brain cDNA library (Stratagene #93206), a human heart cDNA library (Stratagene #936207), and a human aorta cDNA library (Clontech Laboratories #HL1136a, Palo Alto, California). The sequence of the most 5' region of HT9 was utilized to design and synthesize two oligonucleotides, oHT9-1 5'-CCTAGTCCAGTAAACTTGC-3' (SEQ ID NO: 15) and oHT9-4 5'-TTTGCGGCCCTTCCAATATC-3' (SEQ ID NO: 16) which were used to generate a 317 bp PCR probe under conditions described for generating the probe above. While no positive clones were isolated from the heart or aorta cDNA libraries, two positive clones were obtained from the fetal brain library. One of these clones, HFB2, included a cDNA 4.5 Kb insert which included

-21-

approximately 2300 bp of additional sequence. The HFB2 insert corresponds to nucleotides 1 to 3194 of SEQ ID NO: 30.

A composite cDNA encoding $MCCS1\alpha$ was constructed from clones HFB2, HT9 and HT2. The three clones were joined together by digesting HFB2 with the restriction enzymes KpnI and SalI to generate a fragment to comprise the 5' end of the composite clone, digesting HT9 with KpnI and NotI to generate a fragment to comprise the 3' end of the composite clone, and then ligating isolated fragments to the vector pBS SK⁺ (Stratagene) that had been digested with SalI and NotI. The region of the HT9 fragment containing the one nucleotide insertion was replaced with an EcoRV fragment containing nucleotides 3174 to 5282 of clone HT2. The final plasmid containing a 7621 bp insert was named pBSHFB2HT2-27 (ATCC 69951). The DNA and deduced amino acid sequence of the insert are presented in SEQ ID NOs: 1 and 2, respectively. The coding domain of the cDNA initiates with an ATG at nucleotide 333 and ends with a termination codon at nucleotide 7560 predicting a coding sequence of 2409 amino acids and protein of 265 kD. The protein product of the cDNA insert was named $MCCS1\alpha$. Subsequent sequence analysis of the insert in plasmid pBSHFB2HT2-27 (ATCC 69951) revealed sequencing errors in SEQ ID NO: 1. Corrected DNA and deduced amino acid sequences of the insert are set out in SEQ ID NOs: 23 and 24, respectively. Even further sequence analysis of insert in plasmid pBSHFB2HT2-27 revealed sequencing error in SEQ ID NO: 23. At nucleotide position 6317 (SEQ ID NO: 23) a "G" was erroneously included and between positions 6338 and 6339 the sequence was missing an "A". The corrected sequences of $MCCS1\alpha$ are provided in SEQ ID NOs: 30 and 31.

Comparison of the predicted amino acid sequence of $MCCS1\alpha$ with the partial amino acid sequence of $MCCS1\beta$ predicted from clone 517 revealed the presence of a seventy amino acid deletion in the $MCCS1\alpha$ product. The $MCCS1\beta$ clone 517 amino acid sequence corresponds to $MCCS1\alpha$ amino acids 1611 to 2410 of SEQ ID NO: 31. The seventy amino acid deletion in $MCCS1\alpha$ (*i.e.*, where the seventy amino acids would be inserted to generate a product identical to $MCCS1\beta$) occurs between amino acids 2065 and 2066 in SEQ ID NO: 31, seventeen amino acids upstream from the kinase domain. Since both clones maintain an open reading frame, cDNA clone pBSHFB2HT2-27 was apparently generated from alternatively

-22-

spliced mRNA. The carboxyl terminal domains containing the kinase domains are identical in MCCS1 α (amino acids 2083 to 2410 of SEQ ID NO: 31) and MCCS1 β clone (amino acids 543 to 870 of SEQ ID NO: 4).

5 A composite clone containing the complete coding sequence of MCCS1 β (with the seventy amino acid insert) is presented in SEQ ID NO: 32. The amino acid sequence deduced from the clone is presented in SEQ ID NO: 33. This clone is constructed by replacing the sequence between the BSTXI site, which cleaves after nucleotide 3229, and the NotI site in the polylinker sequence at the 3' end of pBSHFB2HT2-27 (SEQ ID NO: 1) with the sequence contained in HT2 between the
10 BstXI site and the NotI site at the 3' end of HT2. Thus this clone contains sequences that are identical to MCCS1 α nucleotides 1 to 5159 of SEQ ID NO: 1 (encoding amino acids 1 to 1609 of SEQ ID NO: 2) linked to sequences that are identical to clone 517 nucleotides 1 to 2610 of SEQ ID NO: 3 (encoding amino acids 1 to 870 of SEQ ID NO: 4). As noted above, subsequent sequence analysis revealed errors
15 in nucleotides 1 to 5159 of SEQ ID NO: 1. Corrected MCCS1 β DNA and deduced amino acid sequences that include the same corrections that appear in MCCS1 α SEQ ID NOs: 23 and 24 are set out in SEQ ID NOs: 25 and 26. The SEQ ID NO: 25 clone represents a cDNA encoding a full length MCCS1 β kinase. Further sequences for MCCS1 β including corrections of errors identified in resequencing the MCCS1 α
20 clone are presented in SEQ ID NOs: 32 and 33.

The MCCS1 products can be divided into three regions based on similarity to other PIK-related kinases: an amino terminal domain (MCCS1 α amino acids 1 to 1081 of SEQ ID NO: 31 and MCCS1 β amino acids 1 to 1150 of SEQ ID NO: 33), a region with similarity to *rad3+* (MCCS1 α amino acids 1082 to 2082 of
25 SEQ ID NO: 31 and MCCS1 β amino acids 1151 to 2151 of SEQ ID NO: 33) and a PIK domain (MCCS1 α amino acids 2083 to 2410 of SEQ ID NO: 31 and MCCS1 β amino acids 2152 to 2480 of SEQ ID NO: 33) including a kinase domain. The amino terminal region and *rad3+* region are regulatory domains that modulate the kinase activity of the enzyme and are involved in interactions with associated proteins.

30 Results of comparisons of the nucleotide and amino acid sequence of MCCS1 α and MCCS1 β to the sequences of other PIK-related and non-PIK-related kinases are shown in Table 1. Specifically, the 3' end of MCCS1 α (nucleotides 6579

-23-

to 7562 of SEQ ID NO: 30 encoding the kinase domain), the 3' end of MCCC1 β (nucleotides 1627 to 2379 of SEQ ID NO: 32 encoding the kinase domain), the *rad3+* domain of MCCC1 α (nucleotides 3576 to 6578 of SEQ ID NO: 30), and the *rad3+* domain of MCCC1 β (clone 517 nucleotides 1 to 1626 of SEQ ID NO: 3) were compared to the analogous region in human ATM [Savitsky *et al.*, *supra*], human DNA-PK [Huntley *et al.*, *Cell*, 82: 849-856 (1995)], human FRAP [Brown *et al.*, *supra*], human p110 [Hu *et al.*, *Mol. Cell. Biol.*, 13(12): 7677-7688 (1993)], *S. cerevisiae* MEC1 [Weinert *et al.*, *Genes Dev.*, 8(6): 652-665 (1994)], *S. pombe* *rad3+* [Seaton *et al.*, *supra* and Hari *et al.*, *Cell*, 82: 815-821 (1995)] and an cAMP-dependent protein kinase (PKA) [Beebe *et al.*, *Mol. Endocrinol.*, 4(3): 465-475 (1990)]. Percent identity of nucleotides is shown in the top line, percent identity of amino acids is shown in the middle line, and percent similarity of amino acids (*i.e.*, including identical amino acids and conservative variations in amino acids) is shown in the bottom line for each kinase in Table 1. Conservative variation as used herein denotes biologically similar residues. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. In the Table, "ND" indicates a value was not determined either because the nucleotide sequence encoding the kinase (*i.e.*, *rad3+*) was not publically available or because the kinase (*i.e.*, FRAP, p110 β , or PKA) lacks the particular domain being compared.

-24-

Table 1

Protein Kinase		MCCS1 α /MCCS1 β Kinase Domain	MCCS1 α <i>rad3</i> + Domain	MCCS1 β <i>rad3</i> + Domain
5	<i>S. pombe rad3</i> +	ND	ND	ND
		56	22	30
		72	46	53
	<i>S. cerevisiae</i> <i>MEC1</i>	51	42	44
		45	21	24
		63	46	49
	Human ATM	50	41	41
		38	22	24
		60	46	47
	Human DNA-PK	43	39	43
		29	19	20
		53	45	49
	Human FRAP	45	ND	ND
		37	ND	ND
		61	ND	ND
10	Human p110 β	45	ND	ND
		24	ND	ND
		54	ND	ND
	Human PKA	39	ND	ND
		16	ND	ND
		39	ND	ND

-25-

Example 2

The MCCS1 gene was mapped to chromosome 3 by a PCR-based assay. Human/rodent somatic cell hybrids containing various human chromosome panels available from the NIGMS Human Genetic Mutant Cell Repository [Drwinga
5 *et al.*, *Genomics*, 16: 311-314 (1993)] were used as templates.

Two oligonucleotide primers oDH23 (SEQ ID NO: 12) and oDH26 5' TGGTTTCTGAGAACATTCCCTGA 3' (SEQ ID NO: 19) based on the MCCS1 α cDNA sequence were utilized to amplify a portion of the gene. The primers generate 237 bp PCR products. PCR conditions consisted of 50 ng genomic DNA, 0.5 μ g of
10 each primer, 200 μ M dNTPs, 1.5mM MgCl₂, 1X PCR buffer (Perkin Elmer-Cetus), and 1 unit of Amplitaq polymerase (Perkin-Elmer Cetus) in a 25 μ l reaction volume. The samples were denatured for 4 minutes and then cycled 35 times with denaturing, annealing, and extension times of 45 seconds, 30 seconds, and 45 seconds, respectively, in a Model 480 Cetus Thermocycler. Five μ l of the resulting PCR
15 product was electrophoresed on a 3% agarose gel and stained with ethidium bromide. DNA corresponding to the human/rodent chromosome 3 hybrid yielded a positive amplification product.

In a second set of amplification reactions, the same oligonucleotide primers were used to sublocalize the MCCS1 gene to a specific region on
20 chromosome 3. The templates for these amplifications consisted of DNA samples from patients with chromosome 3 truncations [Leach *et al.*, *Genomics*, 24: 549-556 (1994)]. Amplifications were performed as described in the foregoing paragraph. The pattern of positive amplification products narrowed the localization to the interval between q21 and q25.1.

Example 3

Polynucleotides encoding carboxyl terminal portions of the PIK-related kinase MCCS1 β were expressed by recombinant techniques in *E. coli*.

Two *E. coli* expression plasmids were constructed that expressed either the COOH-terminal 423 or 571 amino acid residues of the kinase in the Pinpoint
30 fusion protein expression/purification system (Promega, Madison, Wisconsin). Briefly, DNA sequences encoding the COOH-terminal portion of the kinase

-26-

(nucleotides 1339 to 2630 or nucleotides 898 to 2630 of SEQ ID NO: 3) were fused in frame to the COOH-terminus of a 13 kD peptide derived from the transcarboxylase complex from propionibacterium shermanii. This region undergoes biotination in *E. coli*, and thus provides a means for monitoring expression and purification of the fusion proteins. Expression was driven from the tac promoter in pinpoint Xa3. Fusion protein expression was induced with 0.1mM IPTG and confirmed using streptavidin alkaline phosphatase in a pseudo-Western format as described by the manufacturer.

Example 4

Recombinant versions of MCCS1 may also expressed in yeast or in SF9 insect cells using a baculovirus expression system. The FRAP kinase has been expressed, purified and is enzymatically active after expression in the baculovirus system [Brown *et al.*, *supra*].

The coding region of MCCS1 is fused at the amino terminus to a heterologous peptide sequence, such as the FLAG tag MDYKDDDDK (SEQ ID NO: 20) or a six-histidine tag, and reconstructed into the appropriate vectors. Once expressed in insect cells, a monoclonal antibody that recognizes the FLAG tag (Eastman Kodak, Rochester, New York) is used to purify large quantities of the FLAG-PIK-related kinase fusion protein. Infected insect cells are incubated for 48 hours and lysed in lysis buffer (25mM 2-glycerolphosphate, 50mM sodium phosphate pH 7.2, 0.5% Triton-X 100, 2mM EDTA, 2mM EGTA, 25 mM sodium fluoride, 100μM sodium vanadate, 1mM PMSF, 1μg/ml leupeptin, 1μg/ml pepstatin, 1mM benzamidine, and 2mM DTT). Expressed FLAG fusion proteins are purified over a column containing anti-FLAG antibody M2 affinity resin (Eastman Kodak). The column is washed with 20 column volumes of lysis buffer, then 5 column volumes of 0.5M lithium chloride, 50mM Tris pH 7.6, 1mM DTT, and then eluted either with 0.1M glycine pH 3.0 followed by immediate neutralization or by competitive elution with the FLAG peptide. For six-histidine tagged proteins, Ni-NTA agarose (Qiagen) is used for protein purification.

-27-

Shortly after the filing of parent application U.S.S.N. 08/558,666, a gene identified as ATR was described by Antony M. Carr and co-workers (personal communications). ATR appears to encode the same or a closely related protein to MCCS1 based on a comparison of amino acid sequences between ATR and MCCS1. The DNA and deduced amino acid sequences of ATR are presented in SEQ ID NOs: 28 and 29, respectively. The sequence differences between ATR and MCCS1 β are as follows. ATR includes an additional 98 amino acid residues at the N-terminus. At nucleotide position 1284 (SEQ ID NO: 32) there is a conservative base change from "A" in MCCS1 β to "T" in ATR and at nucleotide position 4176, there is an additional conservative base change from "C" in MCCS1 β to "T" in ATR.

The FLAG tag was fused at the amino-terminus of a truncated ATR molecule which lacked the first sixty-six ATR amino acids. The FLAG tag was added by PCR as follows. The oligos FLAG-ATR (5'-CGGGATCCGCCATGGACTACAAGGACGATGACAAGATGTTGCTTGATTTC-3'). And HFB24 (5'-CTTAAGCCGCATGAGCACACCGTC-3') were used in the following PCR reaction: 100ng of pcDNAATR (obtained from Antony M. Carr) as template; 1X PCR buffer (Perkin-Elmer Cetus); 1.5 mM MgCl₂, 200 μ M each of dATP, dGTP, dCTP, and TTP, 10 ng/ μ l of each primer; 1U AmpliTaq (Perkin-Elmer Cetus). The reaction was denatured at 94°C for 4 minutes followed by 30 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. The resulting approximately 800 bp PCR product was digested with BamHI and NheI and was ligated to the 10kb fragment of the mammalian ATR expression plasmid, pcDNAATR digested with BamHI and BstXI along with the remainder of the ATR coding sequences contained on a 2.5 kb BstXI to NheI fragment. Sequence analysis confirmed the addition of the FLAG tag. The insert contained within this plasmid was then used to construct a baculovirus expression plasmid that would express the FLAG tagged ATR truncate. The 5' end of ATR contained on a BamHI to BstXI fragment and the 3' end of ATR contained on a BstXI to SalI fragment derived from pBTM ATR were ligated to the baculovirus expression vector, pFB (Gibco/BRL) that had been digested with BamHI and SalI. This plasmid was designated pFMBCCS β FLAG.

-28-

The full coding region of ATR was fused at the amino terminus to the six histidine tag by PCR. Oligonucleotides MCCS6his (5'-CGGGATCCAGCATGCATCACCATCACCATCACATGGGGGAACATGGGC-3') and Frp1R (5'-CATGACCACTGGCCATTCCACACG-3') were used in a PCR reaction to add the six histidine tag to sequences encoding the amino-terminus of ATR. PCR conditions were as follows: 100 ng of PstA 12ATR (obtained from Antony M. Carr) was used as template; 1X PCR buffer (Perkin-Elmer Cetus); 1.5 mM MgCl₂, 200 μM each of dATP, dGTP, dCTP, and TTP, 10 ng/μl of each primer; 1U AmpliTaq (Perkin-Elmer Cetus). The reaction was denatured at 94°C for 4 minutes followed by 25 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. The approximately 800 bp PCR product was digested with BamHI and MscI and ligated to two other fragments: a 10kb fragment from pcDNAATR digested with BamHI and BstXI and an approximately 3 kb MscI to BstXI fragment containing the remainder of the ATR coding sequence. The addition of the six histidine tag was verified by sequence analysis. The resulting plasmid encoding a six-histidine tagged full length ATR molecule was designated pcDNA6his ATR.

To construct a baculovirus expression plasmid that expressed the entire coding sequence of ATR, the 1.2 kb BamHI to AgeI fragment from pFBMCCSβFLAG was ligated to the BamHI to AgeI fragment from pcDNA6his ATR. The resulting plasmid, designated pFB/HisX6MCCS-1 plasmid was transformed into the E.coli strain, DH5α (Gibco/BRL) for screening of recombinants. This plasmid was purified by using the Promega "Wizard" mini-prep kit, then transformed into *E. coli* αSF9 cells (Invitrogen) using the Cellfectin protocol described by Gibco/BRL.

Forty eight hours after transfection, the SF9 cell pellet and baculovirus produced by the transfected cells were harvested. The virus was stored at 4°C in Grace's Complete media containing 10% FBS, Pennicillin-Streptomycin, and Gentamicin. This viral prep was used to make a high titer (P2) virus stock. The P2 virus stock was used to infect a 50 ml culture of SF9 cells. The cells were collected 48 hours after infection and centrifuged at low speed to pellet the cells without lysis. The cell pellet was stored at -20°C for 24 hours before lysis. The cells were lysed

-29-

in 5 ml of lysis buffer (50 mM Tris, pH 8.0; 500 mM NaCl; 1% NP40; 100 μ M PMSF). Expression of ATR was confirmed by immunoblot using the polyclonal antibody anti-AgDH2 as a probe. The FBHisX6 ATR baculovirus produced an approximately 300 kDa protein that was immunoreactive with anti-AgDH2 antibodies and comigrated with a protein in a mouse testes cell extract.

The P2 virus stock was also used to infect a 2 liter culture of SF9 cells. The cells were collected 48 hours after infection, centrifuged at low speed to pellet the cells without lysis and stored at -20°C. A cell pellet from 150 mls of this culture was lysed in 7.5 ml of lysis buffer (50mM NaPO₄ pH7.2; 0.5% NP-40; 10mM imidazole, 25mM NaF, 100 μ M Na₃VO₄; 0.5mM AEBSF; 1 μ g/ml leupeptin; 1 μ g/ml pepstatin A) and incubated on ice for 15 minutes. The lysate was then centrifuged for 30 minutes at 10,000 x g. The supernatant was removed and any DNA in the lysate resulting from broken nuclei was sheared by aspirating through an 20 gauge needle. Particulate matter was then removed by filtering through a 0.8 micron filter followed by a 0.2 micron filter. This cleared lysate was adjusted to contain 5 mM β -mercaptoethanol and 0.4 M NaCl. A 1 ml Ni-NTA-agarose column (Qiagen) was equilibrated in Buffer A (0.4 M NaCl; 5 mM β -mercaptoethanol; 0.1% Triton X-100; 50 mM NaPO₄ 10 mM imidazole; 25 mM NaF, 100 μ M Na₃VO₄; 0.5 mM AEBSF; 1 μ g/ml leupeptin; 1 μ g/ml pepstatin A) prior to loading the cleared lysate. The sample was loaded at a flow rate of 0.25 ml/minute, washed 5 ml of Buffer A and then eluted in 10 ml of a gradient of 50 to 500 mM imidazole in Buffer A. One half ml fractions were collected and was assayed for kinase activity as follows. Five μ l of each fraction was incubated in kinase buffer, 10 μ Ci ³²P- γ ATP, 10 μ M ATP, and 5 μ g of substrate PHAS-1 (Stratagene) and incubated at 37°C for 20 minutes. The reaction was then spotted onto phosphocellulose spin columns and centrifuged at 2500x g, washed twice with 0.5 ml of 75 mM phosphoric acid and once with 0.5 ml absolute ethanol. The phosphocellulose disks were then transferred to scintillation vials and the counts per minutes incorporated into the PHAS-1 proteins were recorded. Fractions 4 through 9 were found to contain activity toward PHAS-1 and immunoblot analysis confirmed that ATR was also present in the same fractions.

-30-

MCCS1 encoding plasmid DNA was transformed into an *esr1-1* diploid yeast strain (*Mata leu2-1 his4-4 can1 ura3 cyh2 ade6 ade2 esr1-1/MAT a leu2-27 his4 trp1 met2 ade2 esr1-1*), and cells were grown to mid-log phase in either galactose or glucose containing medium. Cells were pelleted, washed and all steps performed at 4°C. Cell pastes were resuspended in buffer (20 mM Tris at pH 8.0, 300 mM NaCl, 10% glycerol, 0.1 mM PMSF, 0.25 mg/ml pepstatin, leupeptin, and aprotinin) and lysed in a French Press or using glass beads. Lysis was verified by microscopy following a low-speed (10K) spin and a high-speed spin (100K), and the supernatant was loaded onto a 1.5 ml Ni-NTA agarose (Qiagen, Inc., Chatsworth, CA) column prewashed in 1x buffer. The column was washed with six column volumes of buffer. The column was eluted stepwise with 8 ml of 10 mM, 50 mM, 100 mM, and 250 mM imidazole in buffer. Fractions were collected and Western analysis was performed using 15 µl of each elution peak. Kinase activity was measured as described above.

15

Example 5

Polyclonal and monoclonal antibodies specific for MCCS1 were generated by standard techniques in the art.

Two different bacterial expression plasmids, pGEX1-MEC and pGEX3-MEC, were constructed for the recombinant production of portions of the MCCS1 polypeptide as fusions to the COOH-terminus of glutathione S-transferase (GST). Both plasmids were used for the generation of antigens AgDH-2 and AgDH-3, from pGEX1-MEC and pGEX3-MEC respectively for use in a standard immunization protocol. pGEX1-MEC contains an EcoRI fragment encoding amino acid residues 566 to 870 of SEQ ID NO: 4 fused to GST in the pGEX1 vector (Pharmacia Biotech, Milwaukee, Wisconsin); pGEX3-MEC contains an Eco RI fragment encoding amino acid residues 118 to 567 of SEQ ID NO: 4 fused to GST in the pGEX3 vector (Pharmacia Biotech). Induction of the pGEX tac promoter with 0.1mM IPTG led to high level expression of each fusion protein in an insoluble form (inclusion bodies). Following lysis of induced cultures with a French pressure cell, AgDH-2 and AgDH-3 extracts were centrifuged through a 35% sucrose solution containing 0.1M NaCl,

-31-

0.01M Tris pH7.5, and 0.001M EDTA (STE). Pellets were then washed twice and resuspended in STE.

For the generation of polyclonal antisera in rabbits, AgDH-2 and AgDH-3 were further purified using preparative SDS polyacrylamide gel electrophoresis and electroelution of each antigen from gel slices. Primary immunization of female New Zealand White rabbits was with 200 μ g of each antigen mixed with complete Freund's adjuvant injected at multiple sites subcutaneously. Subsequent immunizations were with 100 μ g antigen mixed with incomplete Freund's adjuvant at approximately 21 day intervals, and test bleeds were taken after immunizations 3, 4 and 5. Western blot analysis of extracts of human testis tissue demonstrates antibody reactivity against an approximately 270 kD protein in immune but not preimmune antisera. In addition, the immune sera showed reactivity against the M CCS1 pinpoint fusion proteins described in Example 3, providing evidence of the generation of M CCS1-specific antibodies.

The M CCS1-specific antibodies were purified as follows. Inclusion body preparations of AgDH-2 and AgDH-3 were coupled to cyanogen bromide (CNBr)-activated Sepharose (Pharmacia, Alameda, CA). Two mg of antigen were solubilized in 1 % SDS (4.5 ml final volume) and dialyzed overnight against Coupling Buffer (0.1M NaHCO₃/0.1 % SDS). 0.5 ml of 5M NaCl were added to each antigen preparation prior to incubation with the CNBr Sepharose. 0.4 gm of freeze-dried CNBr Sepharose (per antigen) were resuspended in 1 mM HCl and washed in a scintered glass funnel with 250 ml 1 mM HCl added in several aliquots over 15 minutes. The HCl-washed CNBr Sepharose was then removed to a 15 ml snap cap tube and washed twice with 5 ml of Coupling Buffer. Dialyzed antigen preps were added to the washed Sepharose and then incubated at room temperature for 1.5 hours on a slowly rotating wheel. The Sepharose was washed once with 5 ml of Coupling Buffer, once with 10 ml of 0.1M Tris pH8.0, and then incubated in 10 ml 0.1M Tris 8.0 for 2 hours at room temperature to block any remaining reactive groups on the resin. Coupling efficiency was 60-80% as judged by SDS-PAGE analysis. The antigen columns were then washed with 15 ml of 6M Guanidine HCl (to remove uncoupled antigen), 25 ml of Buffer A (50mM Tris pH 7.4), 15 ml of Buffer B (4.5M MgCl₂/1mg/ml BSA/50mM Tris 7.4), and then 50 ml of Buffer A. Thirty ml

-32-

of rabbit serum from immunized animals (rabbit 4747 immunized with AgDH-3 and rabbit 4779 immunized with AgDH-2) were passed over the appropriate antigen column over the course of 3 hours. The columns were then washed with 20 ml of Buffer A, 40 ml of 1M Guanidine HCl, and then equilibrated with an additional 20 ml of Buffer A. Anti-AgDH-3 or Anti-AgDH-2 antibodies were then eluted off the antigen columns with 10 ml of Buffer B. One ml fractions were collected, IgG-containing fractions were pooled and dialyzed against 1 L of phosphate buffered saline (PBS) for 3 hours, and then overnight against 1 L of PBS containing 35 % glycerol.

Antipeptide antibodies were generated against the human ATM protein by coupling a 15-amino-acid peptide (residues 1359-1373) to Keyhole Limpet Hemocyanin-using EDC as described by the manufacturer (Pierce), followed by injection of the coupled immunogen into rabbits. The antibodies were first precipitated from the serum (#6076) with an equal volume of saturated ammonium chloride followed by resuspension and dialysis against PBS. Affinity purification was carried out using a peptide column prepared by coupling the antigenic peptide to CNBr-activated Sepharose (Pharmacia) as described by the manufacturer. The antibodies were then bound to the peptide column and washed with 2 m KCl-PBS. Elution was carried out with 20 ml 5 m NaI (in 1 mM sodium thiosulfate), which was dialyzed immediately against PBS.

To generate monoclonal antibodies, female Balb/c mice were immunized with 50 ug AgDH-2 or AgDH-3. Additional mice were immunized with 25 to 50 ug AgDH-2 or AgDH-3 that had been combined with an equal molar ratio of mAb 61F3B, a monoclonal antibody with specific reactivity to GST. A third group of mice were immunized with SDS polyacrylamide gel slices containing AgDH-2 or AgDH-3. The immunogen for each group of mice was prepared in complete Freund's adjuvant, with subsequent boosts (25 ug antigen in incomplete Freund's) at about 21 day intervals. Cell lines producing monoclonal antibodies were isolated as follows. Briefly a single cell suspension was formed by grinding immunized mouse spleen in serum free RPMI 1640, supplemented with 2mM L-glutamine, 1mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspension was filtered through

-33-

sterile 70-mesh Nitex cell strainer (Becton Dickinson, Parsippany, New Jersey), and washed twice by centrifuging at 200 g for 5 minutes and resuspending the pellet in 20 ml serum free RPMI. Thymocytes taken from three naive Balb/c mice were prepared in this manner.

5 NS-1 myeloma cells kept in log phase in RPMI with 11 % fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, were centrifuged at 200 g for 5 minutes, and the pellet was washed twice as described in the foregoing paragraph. After washing, each cell suspension was brought to a final volume of 10 ml in serum free RPMI, and 10 μ l was diluted
10 10:100. Twenty μ l of each dilution was removed, mixed with 20 μ l 0.4% trypan blue stain in 0.85% saline (Gibco), loaded onto a hemacytometer and counted.

 Two $\times 10^8$ spleen cells were combined with 4 $\times 10^7$ NS-1 cells, centrifuged, and the supernatant was aspirated. The cell pellet was dislodged by tapping the tube and 2 ml of 37°C PEG 1500 (50% in 75mM Hepes, pH 8.0)
15 (Boehringer Mannheim) was added with stirring over the course of 1 minute, followed by adding 14 ml of serum free RPMI over 7 minutes. An additional 16 ml RPMI was added and the cells were centrifuged at 200 g for 10 minutes. After discarding the supernatant, the pellet was resuspended in 200 ml RPMI containing 15% FBS, 100 μ M sodium hypoxanthine, 0.4 μ M aminopterin, 16 μ M thymidine
20 (HAT) (Gibco), 25 units/ml IL-6 (Mallinckrodt, Folcroft, Pennsylvania), and 1.5 $\times 10^6$ thymocytes/ml. The suspension was dispensed into ten 96-well flat bottom tissue culture plates at 200 μ l/well. Cells in plates were fed 3 to 4 times between fusing and screening by aspirating approximately half the medium from each well with an 18 G needle and replenishing plating medium described above except containing 10
25 units/ml IL-6 and lacking thymocytes.

 Fusions were screened when cell growth reached 60-80% confluency (day 7 to 9) by ELISA on AgDH2 versus AgDH3. Immunlon 4 plates (Dynatech, Cambridge, MA) were coated at 4°C overnight with 100 ng/well protein in 30mM carbonate buffer, pH 9.6. Plates were blocked with 100 μ g/well 0.5% fish skin gelatin in PBS for one hour at 37°C, washed 3 times with PBS, 0.05% Tween 20
30 (PBST) and 50 μ l culture supernatant is added. After incubation at 37°C for 30 minutes, and washing as described above, 50 μ l of horseradish peroxidase conjugated

-34-

goat anti-mouse IgG(fc) (Jackson ImmunoResearch, West Grove, PA) diluted 1:10,000 in PBST was added. Plates were incubated as above, washed 4 times with PBST and 100 μ l substrate consisting of 100 μ g/ml of tetramethylbenzidine and 0.15 μ l/ml H_2O_2 in 100mM sodium acetate, pH 5.5, was added. The color reaction was stopped in 5-10 minutes with the addition of 50 μ l of 15% H_2SO_4 . A_{490} was read on a plate reader.

Fifty three pools of hybridomas that were positive in an ELISA were screened for the ability to immunoblot or immunoprecipitate MCCA from a mouse testes cell lysate. Immunoblot analysis using the mouse testes extract is described in Example 6. Immunoprecipitations was performed as follows. A six percent SDS polyacrylamide gel was run and transferred to Immobilon-PVDF in 192 mM glycine, 25 mM Tris base, 0.1% SDS, 20% methanol, then blocked for 1 hour in 5% powdered nonfat milk, 20 mM Tris pH 7.5, 100 mM NaCl 0.1% Tween 20, and cut into the appropriate number of strips. The primary antibody (well supernatant) was diluted in the above block solution and incubated for one hour at room temperature, washed four times in block minus milk, incubated in goat anti-mouse IgG (H+L) HRP (BioRad #170-6516), washed again in block solution minus milk, transferred to NEN Renaissance ECL reagent and developed for 5 minutes.

Immunoprecipitation was performed as follows. Fifty μ l of hybridoma supernatant was incubated for one hour on ice with 300 μ g of testes cell lysate prepared as described in Example 6. Thirty μ l of a 50% slurry of protein A agarose (Pierce, Rockford, IL), prebound to a rabbit anti-mouse bridging antibody (5 μ g/reaction) (Pierce) was added and incubated at 4°C with rocking. The immune complexes were washed three times in lysis buffer and the antigen/antibody complex eluted by boiling in SDS sample buffer (2% SDS, 20 mM Tris pH 6.8, 20% glycerol, 0.001% bromophenol blue). The resulting supernatant was separated on a 6% SDS polyacrylamide gel and transferred to Immobilon-PVDF (Millipore) and an immunoblot was performed using affinity purified rabbit anti-Ag DH2 polyclonal antiserum. Four hybridomas were cloned and characterized in immunoblots, immunoprecipitations and in immunoprecipitation/kinase assays as described in Example 6. The four hybridoma cell lines were designated 224B, 224C (ATCC HB

-35-

12233), 224F (ATCC HB 12234) and 224G. All four monoclonal antibodies recognized M CCS1 by immunoblot and immunoprecipitation.

Example 6

5 M CCS1 associated protein kinase activity was immunoprecipitated using the M CCS1-specific polyclonal antibodies described in Example 5.

Extracts were made from fresh testes tissue isolated from Balb/c mice. Minced testes were homogenized on ice with 10-15 strokes of a tight fitting dounce homogenizer in Lysis Buffer (50 mM NaPO₄, pH 7.2; 0.5% TritonX-100; 2 mM EDTA; 2 mM EGTA; 25 mM NaF; 25 mM 2-glycerophosphate; 1 mM phenylmethylsulfonyl fluoride [PMSF]; 1 µg/ml leupeptin; 1 µg/ml pepstatin A; 2 mM DTT) and incubated on ice for 30 minutes. The lysate was centrifuged at 13,000xg rpm for 10 minutes at 4° C in a TL-100 table-top ultracentrifuge (Beckman) to remove unbroken cells and other insoluble material. Aliquots of cell lysate were snap frozen in liquid N₂ and stored at -70°C. Five hundred ug of testes extract was incubated with either 5 ug of affinity purified anti-AgDH-2 polyclonal antibody or 5 ug purified rabbit IgG (Zymed, So. San Francisco, CA) in 1 ml of Lysis buffer for one hour on ice in microcentrifuge tubes. Thirty µl of protein A sepharose beads (Repligen, Cambridge, MA) (washed in Lysis buffer) were added to the extracts, and then incubated for an additional 30 minutes at 4° C on a rocking platform. The immune complex/Protein A sepharose beads were washed four times with 1 ml of Lysis buffer, one time with 1 ml Kinase Buffer (25 mM Hepes pH 7.7; 50 mM KCl; 10 mM MgCl₂; 0.1% NP-40; 2% glycerol; 1 mM DTT), and then incubated in 20 ul Kinase Buffer with 10 µCi ATP [50 Ci/mmol] for 20 minutes at 37°C. The kinase reactions were stopped with 20 µl 2X SDS sample buffer and heated to 100° C prior to separation on 6% polyacrylamide gels. Gels were fixed in 20% methanol/7% Acetic acid, and then dried onto Whatman 3MM paper prior to autoradiography. While little or no phosphorylation was evident in control immunoprecipitations, immunoprecipitations using anti-AgDH-2 antibody contained two major phosphorylated bands at approximately 300 kD and approximately 180 kD. In addition, there were several minor phosphorylation products, including one which comigrated with the M CCS1 protein itself as demonstrated by Western blot analysis

10
15
20
25
30

-36-

(see Example 8 for Western blot description.) Phosphoaminoacid analysis of the approximately 300 kD protein identified the presence of phosphoserine residues. Addition of 5 ug of AgDH-2 (but not AgDH-3) dramatically reduced or eliminated the MCCS1-associated kinase activity found in the immunoprecipitates.

5

Example 7

The expression pattern of MCCS1 in various human tissues was examined by Northern blot hybridization.

Nylon membranes containing 2 μ g of size-fractionated polyA+ RNA from a variety of human tissue sources were obtained from Clontech Laboratories, Inc., and the hybridization protocol supplied by the manufacturer was followed precisely, except that the final wash was performed at 55° C, rather than 50° C, to minimize the possibility of cross-hybridization to related sequences. The ³²P-labelled DNA hybridization probe used was generated by PCR. A DNA encoding the COOH-terminal 30% of MCCS1 α was used as a template to amplify a 1.3 kb fragment in the presence of ³²P-dCTP using primers 279-3 5'TGGATGATGACAGCTGTGTC 3' (SEQ ID NO: 21) and 279-6 5'TGTAGTCGCTGCTCAATGTC3' (SEQ ID NO: 22).

Results of the Northern blots show that MCCS1 is expressed as an approximately 9 kb mRNA in a wide variety of human tissues. Testis tissue contains the highest level of MCCS1 mRNA, though the transcript is also expressed in small intestine, ovary, prostate, thymus, spleen, heart, peripheral blood lymphocytes, colon, brain, placenta, skeletal muscle, kidney and pancreas.

Expression of MCCS1 mRNA in human cancer cell lines was also examined using a human cancer cell line RNA blot obtained from Clontech. The RNA blot contained RNA from the cell lines HL-60 (promyelocytic leukemia), HeLa (cervical carcinoma), K-562 (chronic myelogenous leukemia), MOLT-4 (lymphoblastic leukemia), Raji (Burkitt's lymphoma), SW480 (colorectal adenocarcinoma), A549 (lung carcinoma), and G361 (melanoma). Northern blot analysis was performed as directed by the manufacturer with hybridization being carried out at 65° C using a 2.0kb KpnI-SalI fragment of the MCCS1 partial clone HFB2. Expression was observed in the HL-60, HeLa, K-562, Raji, SW480, and G361 cell lines with the highest level of expression occurring in the G361 cell line.

-37-

Detectable but low levels of expression were observed in the MOLT-4 and A549 cell lines.

Example 8

5 The expression of MCCA1 mRNA and protein in normal and irradiated mouse testes and in mouse embryos was examined by *in situ* hybridization, immunostaining and/or immunoblotting.

In situ Hybridization

Normal and irradiated mouse testes were harvested from male Balb/c mice. The tissues were sectioned at 6 μ m thickness, picked up on Superfrost Plus[®] (VWR Scientific) slides and allowed to air-dry at room temperature overnight. Sections were
10 stored at -70 $^{\circ}$ C if not immediately used. The tissue sections were fixed in 4% paraformaldehyde (Sigma) in PBS for 20 minutes at 4 $^{\circ}$ C, dehydrated (70%, 95%, 100% ethanol) for 1 minute at 4 $^{\circ}$ C in each grade, then allowed to air dry for 30 minutes at room temperature. The slides were acetylated in a solution of 0.25% (v/v)
15 acetic anhydride (Sigma)/0.1M triethanolamine pH 8.0 for 10 minutes at room temperature with stirring, rinsed in 0.2X SSC for 10 minutes at room temperature with stirring, and dehydrated and air dried as described above. The tissues were hybridized *in situ* with digoxigenin-labeled single-stranded mRNA generated from murine MCCA1 DNA by *in vitro* RNA transcription incorporating digoxigen-UTP
20 (Boehringer Mannheim). The labeled riboprobes (see sequence in SEQ ID NO: 27) (1 μ g/section) and diethylpyrocarbonate (depC)-treated water were added to hybridization buffer with a final concentration of 50% formamide, 0.3 M NaCl, 20 mM Tris pH 7.5, 10% dextran sulfate, 1X Denhardt's solution, 100 mM dithiothreitol (DTT) and 5 mM EDTA, and 20 μ l of the solution was applied to each
25 section and covered with a sterile, RNase-free 22 x 22 cover slip. The mRNA in both the section and the probe solution was denatured by heating the slides to 85 $^{\circ}$ C for 10 minutes in an oven. Hybridization was carried out overnight (12-16 hours) at 50 $^{\circ}$ C.

After hybridization, sections were washed for 1 hour at room
30 temperature in 4X SSC/10 mM DTT, then for 30 minutes at 50 $^{\circ}$ C in 50% formamide/2X SSC/10 mM DTT, 30 minutes at 37 $^{\circ}$ C in a solution of 500 mM

-38-

NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5 (NTE buffer), 30 minutes at 37° C in a bath of 10 µg/mL RNase A (Boehringer Mannheim) in NTE buffer, 15 minutes at 37° C in NTE buffer, 15 minutes at room temperature in 2X SSC, 15 minutes at room temperature in 0.1X SSC, and 2 minutes at room temperature in 100 mM Tris-HCl, 150 mM NaCl, pH 7.5 (Buffer 1). To detect the labeled riboprobes, the sections were blocked for 30 minutes at room temperature in a solution of 5% normal sheep serum (Harlan Bioproducts for Science, Indianapolis, IN) and 0.3% Triton X-100 (Sigma) in Buffer 1 with gentle stirring, after which 150 µl/section of sheep αDigoxigenin-gold conjugate (Goldmark Biologicals, Philipburg, Pa) was applied to the tissues and incubated for 2 hours at room temperature. The slides were then washed three times for 5 minutes in Buffer 1, five times for 3 minutes in sterile deionized water, the excess liquid blotted off the slide and 2 drops each of silver enhancing and initiating solution (Goldmark Biologicals) applied to each section. The chemical reaction was allowed to proceed for 23 minutes at room temperature, then the sections were rinsed thoroughly in sterile deionized water, counterstained in Nuclear Fast Red (Vector), rinsed again in sterile deionized water, air dried overnight at room temperature and mounted with Cytoseal 60 (VWR).

In both normal and irradiated mouse testes signal was observed in the cytoplasm of spermatogonia and spermatocytes. The expression level in irradiated testis was not increased over that seen in normal testis.

Immunostaining

Testis tissue from normal male Balb/c mice was sectioned at 6 µm thickness, picked up on Superfrost Plus® (VWR Scientific) slides and allowed to air-dry at room temperature overnight. Sections were stored at -70° C if not immediately used. The sections were fixed in cold (4° C) acetone for 10 minutes at room temperature; once the slides were removed from the acetone the reagent was allowed to evaporate from the sections. Each tissue section was blocked with 150 µl of a solution of 30% normal rat serum (Harlan Bioproducts), 5% normal goat serum (Vector Laboratories) and 1% bovine serum albumin (BSA) (Sigma) in 1X TBS for 30 minutes at room temperature. After blocking, the solution was gently blotted from the sections and anti-AgDH-3 and anti-AgDH-2 polyclonal antibodies and preimmune sera from the same rabbits were diluted 1:50 and 1:100 in the blocking solution and

-39-

100 μ l applied to each tissue section and incubated for 30 minutes at 37° C. The antibody solution was blotted gently from the sections and unbound antibody removed from the sections by washing the slides 3 times for 5 minutes each in 1X TBS. The excess TBS was blotted from the slide and 100 μ l of the biotinylated goat anti-rabbit antibody contained in the Elite Rabbit IgG Vectastain ABC kit (Vector), prepared according to the product insert, were applied to each section and incubated for 15 minutes at 37° C. After incubation, the slides were washed 2 times in 1X TBS for 5 minutes in each wash. Next, 100 μ l of streptavidin-gold conjugate (Goldmark Biologicals) diluted 1:100 in a solution containing 5% normal rat serum and 1% BSA was applied to each section and incubated for 1 hour at room temperature. The slides were then washed 3 times in 1X TBS for 5 minutes each wash, and 100 μ l of 1% glutaraldehyde (Sigma) in TBS buffer was applied to the slides for 5 minutes at room temperature. The slides were then washed 3 times for 5 minutes each in TBS, then 4 times in sterile deionized water for 3 minutes each. The excess liquid was blotted from each slide and 2 drops each of silver enhancing and initiating solution (Goldmark Biologicals) were applied to each section. The chemical reaction was allowed to proceed for 13 minutes at room temperature, then the sections were rinsed thoroughly in sterile deionized water, counterstained in Nuclear Fast Red (Vector), rinsed again in sterile deionized water, air dried overnight at room temperature and mounted with Cytoseal 60 (VWR).

Signal was detected in the spermatogonia and primary spermatocytes with both of the polyclonal antibodies, but not with the preimmune sera from the same animals.

Immunoblotting

Freshly obtained mouse testicles were minced with razor blades in cold PBS, and a cell suspension was generated using a loose fitting dounce homogenizer. This cell suspension was then boiled with an equal volume of 2X SDS sample buffer. Fifty μ g aliquots of each extract were separated on 6% polyacrylamide gels, transferred onto Immobilon membranes (Millipore, Bedford, MA) and analyzed for anti-MCCS1-reactivity using the affinity purified antibodies in Example 5, and HRP-conjugated goat anti-rabbit secondary antibody and the Renaissance Enhanced Chemiluminescence kit (Dupont/NEN, Boston, MA). Extracts prepared from fresh

-40-

5 mouse testis contain a high molecular weight species (about 294 kD) that was recognized by both affinity-purified antiserum. No reactivity against this protein was seen with either of the preimmune sera. Importantly, the signal obtained from each affinity purified sera was specifically blocked after pre-incubation of the antibody with the corresponding immunogen.

10 In summary, high levels of M CCS1 mRNA and protein are detected in mouse testis in the spermatogonia and primary spermatocytes, cells that are in the early stages of meiosis. This suggests that M CCS1 plays an important role in meiotic cell division. Meiosis is a specialized form of cell division that produces germ cells in higher eukaryotes. There are two major characteristics of meiosis that distinguish it from mitosis. Whereas mitotic cell division results in genetically identical cells containing two of each chromosome, meiotic cell division results in cells containing one of each chromosome. Early in meiosis, during the "reduction division" process, sister chromatids pair and undergo reciprocal recombination at some regions. During 15 this process, these cells are exposed to DNA strand breaks. It is likely that the cellular response to the DNA strand breaks during meiosis is similar to the cellular response found in non-germ cells in response to IR-induced DNA damage. This interpretation is further substantiated by studies that demonstrate the MEC1 is upregulated 10 to 20 fold during sporulation, indicating an important role for M CCS1 20 during meiosis in addition to its role in DNA repair.

Example 9

In order to identify the cells within the developing mouse testis that express M CCS1, Western blot analysis of M CCS1 expression within populations of meiotic cells was performed. Extracts of purified pachytene spermatocytes, round 25 spermatids, condensing spermatids, and epididymal sperm cells were examined for M CCS1 expression as described above in Example 8.

Pachytene spermatocytes, round, and condensing spermatids were prepared from decapsulated testes of adult mice by sequential dissociation with collagenase and trypsin-DNase 1. The cells were separated into discrete populations by sedimentation velocity at unit gravity in 2-4% BSA gradients in Enriched Krebs 30 Ringer Bicarbonate Medium (EKRB). The pachytene spermatocyte and round

-41-

spermatid populations were each at least 85 % pure, while the condensing spermatid population was about 40-50 % pure (contaminated primarily with enucleated residual bodies and some round spermatids). Sperm were obtained from the cauda epididymides. Purified populations of spermatogenic cells were dissolved directly in
5 SDS-sample buffer containing 40 mM DTT, heated to 100° C for 5 minutes, and the amount of protein in each sample determined by the Amido-Black procedure.

The highest levels of MCCC1 protein were found in pachytene spermatocytes, with the level dropping significantly in round spermatids. MCCC1 protein levels were barely detectable lower in the condensing spermatid population,
10 and this may reflect the presence of round spermatids in the preparation (see above). No MCCC1 protein was detected in epididymal sperm. The Western analysis thus corroborates the immunocytochemical data, and suggests a role for MCCC1 in meiotic cells.

Example 10

15 Substrates of MCCC1 and proteins that interact with MCCC1 (for example, members of the cell cycle checkpoint pathway and proteins that localize MCCC1 in cells) may be identified by various assays.

A. Identification of Substrates

Substrates of MCCC1 may be identified by incorporating test
20 compounds in assays for kinase activity. MCCC1 kinase is resuspended in 20 μ l kinase buffer (25mM Hepes pH7.4, 25mM KCl, 10mM MgCl₂, 1mM DTT, 2 % glycerol, 0.1 % NP40, 0.5mM ATP, 10 uCi gamma ³²P-ATP) and incubated for 30 minutes, either in the presence or absence of 4 μ g test compound (*e.g.*, casein, histone H1, or appropriate substrate peptide). Reactions are separated on 12 % PAGE
25 gels and dried onto Whatman paper prior to autoradiography. Moles of phosphate transferred by the kinase to the test compound are measured by autoradiography or scintillation counting. Transfer of phosphate indicates that the test compound is a substrate of the kinase.

The protein PHAS-1 has been identified as an *in vitro* substrate of ATR
30 (Example 4). PHAS-1 is a heat and acid-stable protein that phosphorylated at several sites *in vivo* in response to insulin and growth factors. PHAS-1 binds to the mRNA

-42-

cap binding factor, EIF-4E, and prevents translation of capped mRNAs. Phosphorylation of PHAS-1 at a specific serine residue results in dissociation of PHAS-1 from EIF-4E and thus releasing the inhibition of translation of capped mRNAs. This mechanism allows for a rapid synthesis of protein in response to a particular stimulus. PHAS-1 may be phosphorylated by several protein kinases *in vivo* including a protein kinase that is sensitive to rapamycin. Since the rapamycin-sensitive protein kinase, FRAP, is related to ATR, it would be reasonable to assume that there might be an overlap in substrate specificity between FRAP and ATR and that PHAS-1 is a substrate for both of these protein kinases *in vitro*. To test this hypothesis, ATR that was immunoprecipitated from a mouse testes cell extract or His-tagged ATR purified from baculovirus-infected SF9 cells (Example 4) was incubated with 10 μ g PHAS-1 (Stratagene) in kinase buffer (25 mM Hepes pH 7.4, 25 mM KCl, 10 mM MgCl₂, 1 mM DTT, 0.1% NP-40), 10 μ M ATP and 10 μ Ci₃₂P γ ATP for 20 minutes at 37°C. Since phosphorylated PHAS-1 was known to bind to phosphocellulose paper, the reaction was spotted onto phosphocellulose spin columns and centrifuged at 2500 x g, washed twice with 0.5 ml of 75 mM phosphoric acid and once with 0.5 ml absolute ethanol. The phosphocellulose disks were then transferred to scintillation vials and the counts per minutes incorporated into the PHAS-1 proteins were recorded. ATR readily phosphorylated PHAS-1 whereas negative controls showed little or no PHAS-1 phosphorylation. To map which residue is phosphorylated, the following peptides representing PHAS-1 sequences containing serine and threonine residues were synthesized.

Peptide PH-1

MSGGSSCQTPSRAIPATRR (SEQ ID NO: 36)

Peptide PH-2

GDYSTTPGGTLFSTTPGGTRR (SEQ ID NO: 37)

Peptide PH-3

ECRNSPVTKTRR (SEQ ID NO: 38)

Peptide PH-4

GVTSPSSDEPRR (SEQ ID NO: 39)

Peptide PH-5

MEASQSHLRR (SEQ ID NO: 40)

-43-

Peptide PH-6

RRNSPEDKRAGG (SEQ ID NO: 41)

Peptide PH-7

GEESQFEMDIRR (SEQ ID NO: 42)

5 These peptides are tested in the same kinase reaction to determine which peptide(s) is (are) phosphorylated by ATR. The peptide(s) are then used as substrate for ATR or MCCS1 in assays such as described in Example 11 to identify modulators.

 The same kinase reaction was also used to determine if proteins such as histone H1 (Upstate Biotechnology, Inc., Waltham, NY) and myelin basic protein
10 (Gibco BRL, Gaithersburg, MD) which are known to be substrates of other protein kinases are substrates of MCCS1 and ATR. No phosphorylation of histone H1 or myelin basic protein was observed under the conditions of the assay. Moreover, a peptide from p53 known to be a substrate of DNA-PK was also not phosphorylated in the assay.

15 B. Identification of Interacting Proteins

 Interacting proteins may be identified by the following assays.

 A first assay contemplated by the invention is a two-hybrid screen. The two-hybrid system was developed in yeast [Chien *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582 (1991)] and is based on functional *in vivo* reconstitution of a
20 transcription factor which activates a reporter gene. Specifically, a polynucleotide encoding a protein that interacts with MCCS1 is isolated by: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain; expressing in the host cells a first hybrid
25 DNA sequence encoding a first fusion of part or all of MCCS1 and either the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a library of second hybrid DNA sequences encoding second fusions of part or all of putative MCCS1 binding proteins and the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion;
30 detecting binding of an MCCS1 interacting protein to MCCS1 in a particular host cell by detecting the production of reporter gene product in the host cell; and isolating second hybrid DNA sequences encoding the interacting protein from the particular

-44-

host cell. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

5 Other assays for identifying proteins that interact with M CCS1 may involve immobilizing M CCS1 or a test protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the amount of label bound. Bound label indicates that the test protein interacts with M CCS1.

10 Another type of assay for identifying M CCS1 interacting proteins involves immobilizing M CCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling a test protein with a compound capable of exciting the fluorescent agent, contacting the immobilized M CCS1 with the labelled test protein, detecting light emission by the fluorescent agent, and identifying
15 interacting proteins as test proteins which result in the emission of light by the florescent agent. Alternatively, the putative interacting protein may be immobilized and M CCS1 may be labelled in the assay.

Example 11

Modulators of M CCS1 include M CCS1 variants and other molecules.
20 The modulators may affect M CCS1 kinase activity, its localization in the cell, and/or its interaction with members of the cell cycle checkpoint pathway. Presently preferred regions of M CCS1 which are targets for mutation or the development of selective modulators include the following four regions: the M CCS1 α amino terminal effector domain (amino acids 1 to 1081 of SEQ ID NO: 31), the M CCS1 β amino
25 terminal effector domain (amino acids 1 to 1150 of SEQ ID NO: 33), the M CCS1 α *rad3*+ domain (amino acids 1082 to 2082 of SEQ ID NO: 31), the M CCS1 β *rad3*+ domain (amino acids 1151 to 2151 of SEQ ID NO: 33), the M CCS1 α PIK domain (amino acids 2083 to 2410 of SEQ ID NO: 31), and the M CCS1 β PIK domain (amino acids 2152 to 2480 of SEQ ID NO: 33).

30 M CCS1 variants having mutations in the kinase domain may be useful as a radiosensitizing agents. Mutations specifically contemplated by the invention are,

replacement of the MCCC1 α aspartic acid at amino acid 2241, the asparagine at 2246, and the aspartic acid at 2260 of SEQ ID NO: 31 with alanine or methionine, and the corresponding mutations in MCCC1 β . Analogous mutations in the *rad3+* gene resulted in yeast hypersensitive to radiation. In addition, mutations in the kinase domain of ATM are found in patients with AT, a disease that causes radiation sensitivity.

Furthermore, combinatorial libraries, peptide and peptide mimetics, defined chemical entities, oligonucleotides, and natural product libraries may be screened for activity as modulators in assays such as those described below.

For example, an assay for identifying modulators of MCCC1 kinase activity involves incubating an MCCC1 kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate, both in the presence and absence of a test compound, and measuring the moles of phosphate transferred to the substrate. For example, 2 μ l of the 50 mM imidazole elution pool is added to kinase buffer. (See Example 6.) The reactions are incubated at 37°C for 20 min and samples are analyzed by SDS-PAGE prior to autoradiography or Western analysis.

An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCC1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCC1 kinase.

Moreover, assays for identifying compounds that modulate interaction of MCCC1 with other proteins may involve: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA-binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding a first fusion of part or all of MCCC1 and the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a second hybrid DNA sequence encoding part or all of a protein that interacts with MCCC1 and the DNA binding domain or activating domain of the transcription factor which is not

-46-

incorporated in the first fusion; evaluating the effect of a test compound on the interaction between M CCS1 and the interacting protein by detecting binding of the interacting protein to M CCS1 in a particular host cell by measuring the production of reporter gene product in the host cell in the presence or absence of the test compound; and identifying modulating compounds as those test compounds altering production of the reported gene product in comparison to production of the reporter gene product in the absence of the modulating compound. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

Another type of assay for identifying compounds that modulate the interaction between M CCS1 and an interacting protein involves immobilizing M CCS1 or a natural M CCS1 interacting protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the effect of a test compound on the amount of label bound wherein a reduction in the label bound in the present of the test compound compared to the amount of label bound in the absence of the test compound indicates that the test agent is an inhibitor of M CCS1 interaction with protein. Conversely, an increase in the bound in the presence of the test compound compared to the amount label bound in the absence of the compound indicates that the putative modulator is an activator of M CCS1 interaction with the protein.

Yet another method contemplated by the invention for identifying compounds that modulate the binding between M CCS1 and an interacting protein involves immobilizing M CCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling the interacting protein with a compound capable of exciting the fluorescent agent, contacting the immobilized M CCS1 with the labelled interacting protein in the presence and absence of a test compound, detecting light emission by the fluorescent agent, and identifying modulating compounds as those test compounds that affect the emission of light by the florescent agent in comparison to the emission of light by the fluorescent agent in the absence of the test compound. Alternatively, the M CCS1 interacting protein may be immobilized and M CCS1 may be labelled in the assay.

Example 12

Cell based complementation assays for identifying modulators of M CCS1 or ATM are described below.

5 In one type of assay, host cells (for example, *esr1-1* yeast cells) are transformed with M CCS1-encoding DNA as is described in Example 4. The *esr1-1* yeast strain is normally sensitive to treatment with ultraviolet (UV) light, but *esr1-1* yeast cells expressing M CCS1 or ATR are no longer sensitive to treatment with UV light. The transformed yeast cells are exposed to test compounds and the effect of the test compounds on UV sensitivity of the transformed host cell is determined. Test
10 compounds that are inhibitors of M CCS1 or ATR activity restore UV sensitivity to the M CCS1 transformed *esr1-1* cells. Alternatively, *esr1-1 tell* double mutant yeast cells are used as host cells instead of *esr1-1* yeast cells. The TEL1 gene is homologous to ATM and the TEL1 mutation is described in Morrow, *et al.*, *Cell*, 82:831-840 (1995). The invention also specifically contemplates that the *esr1-1* or
15 *esr1-1 tell* double mutant yeast host cells may be transformed with ATM-encoding DNA (SEQ ID NO: 34).

In an alternative embodiment, the assays include clastogenic agents or events instead of treatment with UV light (*e.g.*, IR, hydroxyurea, or DNA damaging agents). Appropriate host cells for use in such embodiments would be those that are
20 sensitive to the alternative clastogenic agents or events.

Another type of complementation assay involves the use of mammalian host cells such as cell lines derived from cells of AT patients. As described above for yeast cells, the mammalian cells are transfected with DNA encoding M CCS1, ATR, or ATM and then exposed to test compounds. Test compounds that are
25 inhibitors of M CCS1, ATR, or ATM activity will restore the phenotype of the untransformed host cell (*e.g.*, sensitivity to IR).

The above assays can be used to identify compounds that inhibit activity of M CCS1, ATR, and ATM or compounds that inhibit activity of only one of the enzymes.

30 In an alternative type of assay, the yeast or mammalian host cells are transformed with DNA encoding chimeric polypeptides including various combinations of M CCS1 and ATM domains. M CCS1 and ATM show structural

-48-

similarities, and chimeric polypeptides which comprise portions of MCCC1 and ATM are useful in elucidating active sites and binding domains of both MCCC1 and ATM. Polynucleotides encoding the chimeras can be prepared by standard molecular biology techniques known to the skilled worker and as exemplified herein. The chimeric polypeptides are expressed in host cells and modulators of the chimeras can be identified by the assays disclosed herein.

Example 13

MCCC1 and ATM are both involved in meiosis I checkpoints. Since MCCC1 is demonstrated herein to have kinase activity, assays were performed to determine if ATM possessed kinase activity. To determine the kinase activity of ATM, ATM was immunoprecipitated from MRC-5 fibroblasts (ATCC #171-CCL) with polyclonal antisera, 6076. MRC-5 cells are human lung embryonal diploid fibroblasts. MRC-5 cells were obtained from the ATCC at passage 19 and maintained in Minimal Essential Medium supplemented with 10% fetal bovine serum, 100 units/ml penicillin, 100 mg/ml streptomycin, and 100 mM MEM non-essential amino acids. Media and media supplements were obtained through Gibco Life Technologies. Cell lines were maintained in a water-saturated 37°C incubator with 5% C.

MRC-5 cell extracts were prepared by lysis of a 10cm plate of log-phase cells in 0.5 ml of Lysis Buffer I (50 mM NaPO₄, pH 7.2; 0.5% TritonX-100; 2 mM EDTA; 2 mM EGTA; 25 mM NaF; 25 mM 2-glycerophosphate; 1 mM phenylmethylsulfonyl fluoride [PMSF]; 1 µg/ml leupeptin 1 µg/ml pepstatin A; 2 mM DTT) on ice. Cells were scraped from plates using a rubber spatula then sonicated in a cup horn sonicator (Sonifier 250, Branson Ultrasonics Corp., Danbury, CT) at 100% output for 90 seconds. Lysates were then clarified in a 4°C microfuge for 2 minutes. Preclearing was done by adding 10 µg purified rabbit IgG (Zymed) and 30 µl Protein A Agarose slurry (Pierce) followed by incubation at 4°C for 60 minutes while rocking. To the precleared lysates, 10 µg of affinity purified 6076 antisera (or 10 µg 6076 pre-blocked with 0.04 mg P45 peptide for 30 min.) was added and incubated on ice for 60 minutes. Immunoprecipitates were collected by addition of

-49-

30 μ l Protein A agarose slurry and incubated with rocking at 4°C for 30 minutes followed by four washes in Lysis Buffer I.

Kinase reactions were carried out by washing the immunoprecipitations once with kinase buffer (25 μ M Hepes pH 7.7; 50 mM KCl; 10 mM MgCl₂; 0.1 % NP-40; 2 % glycerol; 1 mM DTT), followed by incubation in 20 μ l of Kinase Buffer containing 10 μ M ATP + 10 μ Ci γ ³²P-ATP [50 Ci/mmol] for 20 minutes at 37°C. Reactions were stopped by the addition of 20 μ l 2X SDS sample buffer and boiled for 5 minutes prior to separation on 6% SDS polyacrylamide gels. The gels were dried and exposed to x-ray film (Kodak, XAR-5) at -80°C overnight.

10 10 cm plates of log-phase MRC-5 cells were washed once with PBS then incubated in Dulbecco's Modified Eagle Medium (minus methionine) containing 2% dialyzed fetal bovine serum for 30 minutes. Cells were labeled by adding 200 μ Ci ³⁵S-methionine (1000 Ci/mmol TRAN³⁵S-LABEL, ICN Radiochemicals) for 2 hours. Labeled cells were then washed once with PBS and frozen at -80°C prior to immunoprecipitation.

The incubation of the immunoprecipitated complexes in kinase buffer produced a phosphorylated product with a molecular weight of approximately 350,000 that co-migrated with ATM in polyacrylamide gels.

Similar results were obtained for ATR immune complexes immunoprecipitated with anti-AgDH-2 (MCCS1) polyclonal antisera of Example 5. ATR and ATM thus appear to be able to self-phosphorylate or associate with a protein kinase.

To determine the role of ATR and ATM in meiosis, immunostaining techniques on surface spreads of mouse spermatocytes were utilized to localize ATR and ATM to meiotic chromosomes. Antibodies recognizing ATR and ATM were utilized with mouse antibodies against Cor1. Cor1 is a component of axial/lateral elements of synapsing chromosomes [Dobson *et al.*, *J. Cell Sci.*, 107:2749-2760 (1994)]. Cor1 chromosomal staining appears when the axial elements begin to form between the sister chromatids of each homolog in leptotema of meiotic prophase, prior to the initiation of synapsis. As homologous bivalents synapse, the axial elements from the two homologs align and a central element forms between them, completing the structure called the synaptonemal complex (SC).

-50-

When short stretches of Cor1 begin to appear prior to any evidence of synapsis, neither ATR nor ATM is detectable. As homologs start synapsis, both proteins were seen at pairing forks; however, the location and behavior of the two proteins differed markedly. In normal zygotene nuclei, the stage during which homologs synapse, ATR was present in small amounts and transiently at discrete foci along the asynapsed (unpaired) axes. As homologs synapse, ATR disappeared from these locations. However, at regions delayed in synapsis, often seen near the proximal ends of autosomal bivalents, there was an accumulation of ATR foci along the unsynapsed axes. ATR was detected at similar locations on the two axial elements. In nuclei where an entire autosome fails to find its homologous pairing partner, ATR foci were detected along the entire lengths of these asynapsed axis. In males, where the X chromosome has no homolog, ATR foci were localized along the unpaired axis.

ATM was also visualized as foci and was first detected during zygonema as homologs synapse, but ATM localization was different than ATR. ATM was first observed along synapsed axes when homologous autosomal axial elements come into contact. However, during mid-pachynema, after autosomal synapsis has been completed, ATM foci appeared on the X chromosome axis. ATM localization persisted on fully synapsed bivalents into pachynema, a substage that lasts 3 days in mouse oocytes and 6 days in mouse spermatocytes. During pachynema, the number of foci drops gradually, stabilizing briefly in mid-pachynema before eventually disappearing mid-to late pachynema. Thus, ATR and ATM protein kinases play important and complementary roles at distinct stages in meiosis I.

The involvement of ATR appears to be transient during early meiotic prophase while the role of ATM appears to be more prolonged. However, both ATR and ATM coordinate the various events of meiotic prophase by performing similar checkpoint functions.

The foregoing illustrative examples relate to presently preferred embodiments of the invention and numerous modifications and variations thereof are expected to occur to those skilled in the art. Thus only such limitations as appear in the appended claims should be placed upon the scope of the present invention.

-51-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: ICOS Corporation
- (ii) TITLE OF INVENTION: Cell Cycle Checkpoint PIK-Related Kinase Materials and Methods
- (iii) NUMBER OF SEQUENCES: 42
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 - (B) STREET: 6300 Sears Tower, 233 S. Wacker Dr.
 - (C) CITY: Chicago
 - (D) STATE: Illinois
 - (E) COUNTRY: USA
 - (F) ZIP: 60606
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Noland, Greta E.
 - (B) REGISTRATION NUMBER: 35,302
 - (C) REFERENCE/DOCKET NUMBER: 27866/33607
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (312) 474-6300
 - (B) TELEFAX: (312) 474-0448
 - (C) TELEX: 25-3856

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7621 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pBSHFB2HT2-27
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 333..7559
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

-52-

CTTGTGAAGA GAATGTTTTA CACTCTTGTT AGTGAAGTTT ATTCTTTAAA AGTCAATCGT	60
CAAGGATTTA GCAAATGAAT TAGCACTTCG GATATACTTG TTTATTTAAT ATCTTTTTTG	120
TTTATTTCAA AGAATTCAGT AATTGGATCA TAACGAGACT TCTGCGGATT GCAGCAACTC	180
CCTCCTGTCA TTTGTTACAC AAGAAAATCT GTGAAGTCAT CTGTTTCATTA TTATTTCTTT	240
TTAAAAGCAA GAGTCCTGCT ATTTTGGGG TACTCACAAA AGAATTATTA CAACTTTTTG	300
AAGACTTGGT TTACCTCCAT AGAAGAAATG TG ATG GGT CAT GCT GTG GAA TGG	353
Met Gly His Ala Val Glu Trp	
1 5	
CCA GTG GTC ATG AGC CGA TTT TTA AGT CAA TTA GAT GAA CAC ATG GGA	401
Pro Val Val Met Ser Arg Phe Leu Ser Gln Leu Asp Glu His Met Gly	
10 15 20	
TAT TTA CAA TCA GCT CCT TTG CAG TTG ATG AGT ATG CAA AAA TTA GAA	449
Tyr Leu Gln Ser Ala Pro Leu Gln Leu Met Ser Met Gln Lys Leu Glu	
25 30 35	
TTT ATT GAA GTC ACT TTA TTA ACG GTT CTT ACT CGT ATT ATT GCA ATT	497
Phe Ile Glu Val Thr Leu Leu Thr Val Leu Thr Arg Ile Ile Ala Ile	
40 45 50 55	
GTG TTT TTT AGA AGG CAA GAA CTC TTA CTT TGG CAG ATA GGT TGT GTT	545
Val Phe Phe Arg Arg Gln Glu Leu Leu Leu Trp Gln Ile Gly Cys Val	
60 65 70	
CTG CTA GAG TAT GGT AGT CCA AAA ATT AAA TCC CTA GCA ATT AGC TTT	593
Leu Leu Glu Tyr Gly Ser Pro Lys Ile Lys Ser Leu Ala Ile Ser Phe	
75 80 85	
TTA ACA GAA CTT TTT CAG CTT GGA GGA CTA CCA GCA CAA CCA GCT AGC	641
Leu Thr Glu Leu Phe Gln Leu Gly Gly Leu Pro Ala Gln Pro Ala Ser	
90 95 100	
ACT TTT TTC AGC TCA TTT TTG GAA TTA TTA AAA CAC CTT GTA GAA ATG	689
Thr Phe Phe Ser Ser Phe Leu Glu Leu Leu Lys His Leu Val Glu Met	
105 110 115	
GAT ACT GAC CAA TTG AAA CTC TAT GAA GAG CCA TTA TCA AAG CTG ATA	737
Asp Thr Asp Gln Leu Lys Leu Tyr Glu Glu Pro Leu Ser Lys Leu Ile	
120 125 130 135	
AAG ACA CTA TTT CCC TTT GAA GCA GAA GCT TAT AGA AAT ATT GAA CCT	785
Lys Thr Leu Phe Pro Phe Glu Ala Glu Ala Tyr Arg Asn Ile Glu Pro	
140 145 150	
GTC TAT TTA AAT ATG CTG CTG GAA AAA CTC TGT GTC ATG TTT GAA GAC	833
Val Tyr Leu Asn Met Leu Leu Glu Lys Leu Cys Val Met Phe Glu Asp	
155 160 165	
GGT GTG CTC ATG CGG CTT AAG TCT GAT TTG CTA AAA GCA GCT TTG TGC	881
Gly Val Leu Met Arg Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys	
170 175 180	
CAT TTA CTG CAG TAT TTC CTT AAA TTT GTG CCA GCT GGG TAT GAA TCT	929
His Leu Leu Gln Tyr Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser	
185 190 195	
GCT TTA CAA GTC AGG AAG GTC TAT GTG AGA AAT ATT TGT AAA GCT CTT	977
Ala Leu Gln Val Arg Lys Val Tyr Val Arg Asn Ile Cys Lys Ala Leu	
200 205 210 215	

-53-

TTG GAT GTG CTT GGA ATT GAG GTA GAT GCA GAG TAC TTG TTG GGC CCA Leu Asp Val Leu Gly Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro 220 225 230	1025
CTT TAT GCA GCT TTG AAA ATG GAA AGT ATG GAA ATC ATT GAG GAG ATT Leu Tyr Ala Ala Leu Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile 235 240 245	1073
CAA TGC CAA ACT CAA CAG GAA AAC CTC AGC AGT AAT AGT GAT GGA ATA Gln Cys Gln Thr Gln Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile 250 255 260	1121
TCA CCC AAA AGG CGT CGT CTC AGC TCG TCT CTA AAC CCT TCT AAA AGA Ser Pro Lys Arg Arg Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg 265 270 275	1169
GCA CCA AAA CAG ACT GAG GAA ATT AAA CAT GTG GAC ATG AAC CAA AAG Ala Pro Lys Gln Thr Glu Glu Ile Lys His Val Asp Met Asn Gln Lys 280 285 290 295	1217
AGC ATA TTA TGG AGT GCA CTG AAA CAG AAA GCT GAA TCC CTT CAG ATT Ser Ile Leu Trp Ser Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile 300 305 310	1265
TCC CTT GAA TAC AGT GGC CTA AAG AAT CCT GTT ATT GAG ATG TTA GAA Ser Leu Glu Tyr Ser Gly Leu Lys Asn Pro Val Ile Glu Met Leu Glu 315 320 325	1313
GGA ATT GCT GTT GTC TTA CAA CTG ACT GCT CTG TGT ACT GTT CAT TGT Gly Ile Ala Val Val Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys 330 335 340	1361
TCT CAT CAA AAC ATG AAC TGC CGT ACT TTC AAG GAC TGT CAA CAT AAA Ser His Gln Asn Met Asn Cys Arg Thr Phe Lys Asp Cys Gln His Lys 345 350 355	1409
TCC AAG AAG AAA CCT TCT GTA GTG ATA ACT TGG ATG TCA TTG GAT TTT Ser Lys Lys Lys Pro Ser Val Val Ile Thr Trp Met Ser Leu Asp Phe 360 365 370 375	1457
TAC ACA ACA GTG CTT AAG AGC TGT AGA AGG TTG TTA GAA TCT GTT CAG Tyr Thr Thr Val Leu Lys Ser Cys Arg Arg Leu Leu Glu Ser Val Gln 380 385 390	1505
AAA CGG ACT GGA GGC AAC ATT GAT AAG GTG GTG AAA ATT TAT GAT GCT Lys Arg Thr Gly Gly Asn Ile Asp Lys Val Val Lys Ile Tyr Asp Ala 395 400 405	1553
TTG ATT TAT ATG CAA GTA AAC AGT TCA TTT GAA GAT CAT ATC CTG GAA Leu Ile Tyr Met Gln Val Asn Ser Ser Phe Glu Asp His Ile Leu Glu 410 415 420	1601
GAT TTA TGT GGA ATG CTC TCA CTT CCA TGG ATT TAT TCC CAT TCT GAT Asp Leu Cys Gly Met Leu Ser Leu Pro Trp Ile Tyr Ser His Ser Asp 425 430 435	1649
GAT GGC TGT TTA AAG TTG ACC ACA TTT GCC GCT AAT CTT CTA ACA TTA Asp Gly Cys Leu Lys Leu Thr Thr Phe Ala Ala Asn Leu Leu Thr Leu 440 445 450 455	1697
AGC TGT AGG ATT TCA GAT AGC TAT TCA CCA CAG GCA CAA TCA CGA TGT Ser Cys Arg Ile Ser Asp Ser Tyr Ser Pro Gln Ala Gln Ser Arg Cys 460 465 470	1745
GTG TTT CTT CTG ACT CTG TTT CCA AGA AGA ATA TTC CTT GAG TGG AGA Val Phe Leu Leu Thr Leu Phe Pro Arg Arg Ile Phe Leu Glu Trp Arg 475 480 485	1793

-54-

ACA GCA GTT TAC AAC TGG GCC CTG CAG AGC TCC CAT GAA GTA ATC CGG Thr Ala Val Tyr Asn Trp Ala Leu Gln Ser Ser His Glu Val Ile Arg 490 495 500	1841
GCT AGT TGT GTT AGT GGA TTT TTT ATC TTA TTG CAG CAG CAG AAT TCT Ala Ser Cys Val Ser Gly Phe Phe Ile Leu Leu Gln Gln Gln Asn Ser 505 510 515	1889
TGT AAC AGA GTT CCC AAG ATT CTT ATA GAT AAA GTC AAA GAT GAT TCT Cys Asn Arg Val Pro Lys Ile Leu Ile Asp Lys Val Lys Asp Asp Ser 520 525 530 535	1937
GAC ATT GTC AAG AAA GAA TTT GCT TCT ATA CTT GGT CAA CTT GTC TGT Asp Ile Val Lys Lys Glu Phe Ala Ser Ile Leu Gly Gln Leu Val Cys 540 545 550	1985
ACT CTT CAC GGC ATG TTT TAT CTG ACA AGT TCT TTA ACA GAA CCT TTC Thr Leu His Gly Met Phe Tyr Leu Thr Ser Ser Leu Thr Glu Pro Phe 555 560 565	2033
TCT GAA CAC GGA CAT GTG GAC CTC TTC TGT AGG AAC TTG AAA GCC ACT Ser Glu His Gly His Val Asp Leu Phe Cys Arg Asn Leu Lys Ala Thr 570 575 580	2081
TCT CAA CAT GAA TGT TCA TCT TCT CAA CTA AAA GCT TCT GTC TGC AAG Ser Gln His Glu Cys Ser Ser Ser Gln Leu Lys Ala Ser Val Cys Lys 585 590 595	2129
CCA TTC CTT TTC CTA CTG AAA AAA AAA ATA CCT AGT CCA GTA AAA CTT Pro Phe Leu Phe Leu Leu Lys Lys Lys Ile Pro Ser Pro Val Lys Leu 600 605 610 615	2177
GCT TTC ATA GAT AAT CTA CAT CAT CTT TGT AAG CAT CTT GAT TTT AGA Ala Phe Ile Asp Asn Leu His His Leu Cys Lys His Leu Asp Phe Arg 620 625 630	2225
GAA GAT GAA ACA GAT GTA AAA GCA GTT CTT GGA ACT TTA TTA AAT TTA Glu Asp Glu Thr Asp Val Lys Ala Val Leu Leu Gly Thr Leu Leu Asn Leu 635 640 645	2273
ATG GAA GAT CCA GAC AAA GAT GTT AGA GTG GCT TTT AGT GGA AAT ATC Met Glu Asp Pro Asp Lys Asp Val Arg Val Ala Phe Ser Gly Asn Ile 650 655 660	2321
AAG CAC ATA TTG GAA TCC TTG GAC TCT GAA GAT GGA TTT ATA AAG GAG Lys His Ile Leu Glu Ser Leu Asp Ser Glu Asp Gly Phe Ile Lys Glu 665 670 675	2369
CTT TTT GTC TTA AGA ATG AAG GAA GCA TAT ACA CAT GCC CAA ATA TCA Leu Phe Val Leu Arg Met Lys Glu Ala Tyr Thr His Ala Gln Ile Ser 680 685 690 695	2417
AGA AAT AAT GAG CTG AAG GAT ACC TTG ATT CTT ACA ACA GGG GAT ATT Arg Asn Asn Glu Leu Lys Asp Thr Leu Ile Leu Thr Thr Gly Asp Ile 700 705 710	2465
GGA AGG GCC GCA AAA GGA GAT TTG GTA CCA TTT GCA CTC TTA CAC TTA Gly Arg Ala Ala Lys Gly Asp Leu Val Pro Phe Ala Leu Leu His Leu 715 720 725	2513
TTG CAT TGT TTG TTA TCC AAG TCA GCA TCT GTC TCT GGA GCA GCA TAC Leu His Cys Leu Leu Ser Lys Ser Ala Ser Val Ser Gly Ala Ala Tyr 730 735 740	2561
ACA GAA ATT AGA GCT CTG GTT GCA GCT AAA AGT GTT AAA CTG CAA AGT Thr Glu Ile Arg Ala Leu Val Ala Ala Lys Ser Val Lys Leu Gln Ser 745 750 755	2609

-55-

TTT TTC AGC CAG TAT AAG AAA CCC ATC TGT CAG TTT TTG GTA GAA TCC Phe Phe Ser Gln Tyr Lys Lys Pro Ile Cys Gln Phe Leu Val Glu Ser 760 765 770 775	2657
CTT CAC TCT AGT CAG ATG ACA GCA CTT CCG AAT ACT CCA TGC CAG AAT Leu His Ser Ser Gln Met Thr Ala Leu Pro Asn Thr Pro Cys Gln Asn 780 785 790	2705
GCT GAC GTG CGA AAA CAA GAT GTG GCT CAC CAG AGA GAA ATG GCT TTA Ala Asp Val Arg Lys Gln Asp Val Ala His Gln Arg Glu Met Ala Leu 795 800 805	2753
AAT ACG TTG TCT GAA ATT GCC AAC GTT TTC GAC TTT CCT GAT CTT AAT Asn Thr Leu Ser Glu Ile Ala Asn Val Phe Asp Phe Pro Asp Leu Asn 810 815 820	2801
CGT TTT CTT ACT AGG ACA TTA CAA GTT CTA CTA CCT GAT CTT GCT GCC Arg Phe Leu Thr Arg Thr Leu Gln Val Leu Leu Pro Asp Leu Ala Ala 825 830 835	2849
AAA GCA AGC CCT GCA GCT TCT GCT CTC ATT CGA ACT TTA GGA AAA CAA Lys Ala Ser Pro Ala Ala Ser Ala Leu Ile Arg Thr Leu Gly Lys Gln 840 845 850 855	2897
TTA AAT GTC AAT CGT AGA GAG ATT TTA ATA AAC AAC TTC AAA TAT ATT Leu Asn Val Asn Arg Arg Glu Ile Leu Ile Asn Asn Phe Lys Tyr Ile 860 865 870	2945
TTT TCT CAT TTG GTC TGT TCT TGT TCC AAA GAT GAA TTA GAA CGT GCC Phe Ser His Leu Val Cys Ser Cys Ser Lys Asp Glu Leu Glu Arg Ala 875 880 885	2993
CTT CAT TAT CTG AAG AAT GAA ACA GAA ATT GAA CTG GGG AGC CTG TTG Leu His Tyr Leu Lys Asn Glu Thr Glu Ile Glu Leu Gly Ser Leu Leu 890 895 900	3041
AGA CAA GAT TTC CAA GGA TTG CAT AAT GAA TTA TTG CTG CGT ATT GGA Arg Gln Asp Phe Gln Gly Leu His Asn Glu Leu Leu Leu Arg Ile Gly 905 910 915	3089
GAA CAC TAT CAA CAG GTT TTT AAT GGT TTG TCA ATA CTT GCC TCA TTT Glu His Tyr Gln Gln Val Phe Asn Gly Leu Ser Ile Leu Ala Ser Phe 920 925 930 935	3137
GCA TCC AGT GAT GAT CCA TAT CAG GGC CCG AGA GAT ATC ATA TCA CCT Ala Ser Ser Asp Asp Pro Tyr Gln Gly Pro Arg Asp Ile Ile Ser Pro 940 945 950	3185
GAA CTG ATG GCT GAT TAT TTA CAA CCC AAA TTG TTG GGC ATT TTG GCT Glu Leu Met Ala Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu Ala 955 960 965	3233
TTT TTT AAC ATG CAG TTA CTG AGC TCT AGT GTT GGC ATT GAA GAT AAG Phe Phe Asn Met Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp Lys 970 975 980	3281
AAA ATG GCC TTG AAC AGT TTG ATG TCT TTG ATG AAG TTA ATG GGA CCC Lys Met Ala Leu Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly Pro 985 990 995	3329
AAA CAT GTC AGT TCT GTG AGG GTG AAG ATG ATG ACC ACA CTG AGA ACT Lys His Val Ser Ser Val Arg Val Lys Met Met Thr Thr Leu Arg Thr 1000 1005 1010 1015	3377
GGC CTT CGA TTC AAG GAT GAT TTT CCT GAA TTG TGT TGC AGA GCT TGG Gly Leu Arg Phe Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala Trp 1020 1025 1030	3425

-56-

GAC TGC TTT GTT CGC TGC CTG GAT CAT GCT TGT CTG GGC TCC CTT CTC Asp Cys Phe Val Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu Leu 1035 1040 1045	3473
AGT CAT GTA ATA GTA GCT TTG TTA CCT CTT ATA CAC ATC CAG CCT AAA Ser His Val Ile Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro Lys 1050 1055 1060	3521
GAA ACT GCA GCT ATC TTC CAC TAC CTC ATA ATT GAA AAC AGG GAT GCT Glu Thr Ala Ala Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp Ala 1065 1070 1075	3569
GTG CAA GAT TTT CTT CAT GAA ATA TAT TTT TTA CCT GAT CAT CCA GAA Val Gln Asp Phe Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro Glu 1080 1085 1090 1095	3617
TTA AAA AAG ATA AAA GCC GTT CTC CAG GAA TAC AGA AAG GAG ACC TCT Leu Lys Lys Ile Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr Ser 1100 1105 1110	3665
GAG AGC ACT GAT CTT CAG ACA ACT CTT CAG CTC TCT ATG AAG GCC ATT Glu Ser Thr Asp Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala Ile 1115 1120 1125	3713
CAA CAT GAA AAT GTC GAT GTT CGT ATT CAT GCT CTT ACA AGC TTG AAG Gln His Glu Asn Val Asp Val Arg Ile His Ala Leu Thr Ser Leu Lys 1130 1135 1140	3761
GAA ACC TTG TAT AAA AAT CAG GAA AAA CTG ATA AAG TAT GCA ACA GAC Glu Thr Leu Tyr Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr Asp 1145 1150 1155	3809
AGT GAA ACA GTA GAA CCT ATT ATC TCA CAG TTG GTG ACA GTG CTT TTG Ser Glu Thr Val Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu Leu 1160 1165 1170 1175	3857
AAA GGT TGC CAA GAT GCA AAC TCT CAA GCT CGG TTG CTC TGT GGG GAA Lys Gly Cys Gln Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly Glu 1180 1185 1190	3905
TGT TTA GGG GAA TTG GGG GCG ATA GAT CCA GGT CGA TTA GAT TTC TCA Cys Leu Gly Glu Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe Ser 1195 1200 1205	3953
ACA ACT GAA ACT CAA GGA AAA GAT TTT ACA TTT GTG ACT GGA GTA GAA Thr Thr Glu Thr Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val Glu 1210 1215 1220	4001
GAT TCA AGC TTT GCC TAT GGA TTA TTG ATG GAG CTA ACA AGA GCT TAC Asp Ser Ser Phe Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala Tyr 1225 1230 1235	4049
CTT GCG TAT GCT GAT AAT AGC CGA GCT CCA GAT TCA GCT GCC TAT GCC Leu Ala Tyr Ala Asp Asn Ser Arg Ala Pro Asp Ser Ala Ala Tyr Ala 1240 1245 1250 1255	4097
ATT CAG GAG TTG CTT TCT ATT TAT GAC TGT AGA GAG ATG GAG ACC AAC Ile Gln Glu Leu Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr Asn 1260 1265 1270	4145
GGC CCA GGT CAC CAA TTG TGG AGG AGA TTT CCT GAG CAT GTT CGG GAA Gly Pro Gly His Leu Leu Trp Arg Arg Phe Pro Glu His Val Arg Glu 1275 1280 1285	4193
ATA CTA GAA CCT CAT CTA AAT ACC AGA TAC AAG AGT TCT CAG AAG TCA Ile Leu Glu Pro His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys Ser 1290 1295 1300	4241

-57-

ACC GAT TGG TCT GGA GTA AAG AAG CCA ATT TAC TTA AGT AAA TTG GGT Thr Asp Trp Ser Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu Gly 1305 1310 1315	4289
AGT AAC TTT GCA GAA TGG TCA GCA TCT TGG GCA GGT TAT CTT ATT ACA Ser Asn Phe Ala Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile Thr 1320 1325 1330 1335	4337
AAG GTT CGA CAT GAT CTT GCC AGT AAA ATT TTC ACC TGC TGT AGC ATT Lys Val Arg His Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser Ile 1340 1345 1350	4385
ATG ATG AAG CAT GAT TTC AAA GTG ACC ATC TAT CTT CTT CCA CAT ATT Met Met Lys His Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His Ile 1355 1360 1365	4433
CTG GTG TAT GTC TTA CTG GGT TGT AAT CAA GAA GAT CAG CAG GAG GTT Leu Val Tyr Val Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu Val 1370 1375 1380	4481
TAT GCA GAA ATT ATG GCA GTT CTA AAG CAT GAC GAT CAG CAT ACC ATA Tyr Ala Glu Ile Met Ala Val Leu Lys His Asp Asp Gln His Thr Ile 1385 1390 1395	4529
AAT ACC CAA GAC ATT GCA TCT GAT CTG TGT CAA CTC AGT ACA CAG ACT Asn Thr Gln Asp Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln Thr 1400 1405 1410 1415	4577
GTG TTC TCC ATG CTT GAC CAT CTC ACA CAG TGG GCA AGG CAC AAA TTT Val Phe Ser Met Leu Asp His Leu Thr Gln Trp Ala Arg His Lys Phe 1420 1425 1430	4625
CAG GCA CTG AAA GCT GAG AAA TGT CCA CAC AGC AAA TCA AAC AGA AAT Gln Ala Leu Lys Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg Asn 1435 1440 1445	4673
AAG GTA GAC TCA ATG GTA TCT ACT GTG GAT TAT GAA GAC TAT CAG AGT Lys Val Asp Ser Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln Ser 1450 1455 1460	4721
GTA ACC CGT TTT CTA GAC CTC ATA CCC CAG GAT ACT CTG GCA GTA GCT Val Thr Arg Phe Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val Ala 1465 1470 1475	4769
TCC TTT CGC TCC AAA GCA TAC ACA CGA GCT GTA ATG CAC TTT GAA TCA Ser Phe Arg Ser Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu Ser 1480 1485 1490 1495	4817
TTT ATT ACA GAA AAG AAG CAA AAT ATT CAG GAA CAT CTT GGA TTT TTA Phe Ile Thr Glu Lys Lys Gln Asn Ile Gln Glu His Leu Gly Phe Leu 1500 1505 1510	4865
CAG AAA TTG TAT GCT GCT ATG CAT GAA CCT GAT GGA GTG TCC GGA GTC Gln Lys Leu Tyr Ala Ala Met His Glu Pro Asp Gly Val Ser Gly Val 1515 1520 1525	4913
AGT GCA ATT AGA AAG GCA GAA CCA TCT CTA AAA GAA CAG ATC CTT GAA Ser Ala Ile Arg Lys Ala Glu Pro Ser Leu Lys Glu Gln Ile Leu Glu 1530 1535 1540	4961
CAT GAA AGC CTT GGC TTG CTG AGG GAT GCC ACT GCT TGT TAT GAC AGG His Glu Ser Leu Gly Leu Leu Arg Asp Ala Thr Ala Cys Tyr Asp Arg 1545 1550 1555	5009
GCT ATT CAG CTA GAA CCA GAC CAG ATC ATT CAT TAC CAT GGT GTA GTA Ala Ile Gln Leu Glu Pro Asp Gln Ile Ile His Tyr His Gly Val Val 1560 1565 1570 1575	5057

-58-

AAG TCC ATG TTA GGT CTT GGT CAG CTG TCT ACT GTT ATC ACT CAG GTG Lys Ser Met Leu Gly Leu Gly Gln Leu Ser Thr Val Ile Thr Gln Val 1580 1585 1590	5105
AAT GGA GTG CAT GCT AAC AGG TCC GAG TGG ACA GAT GAA TTA AAC ACG Asn Gly Val His Ala Asn Arg Ser Glu Trp Thr Asp Glu Leu Asn Thr 1595 1600 1605	5153
TAC AGA GTG GAA GCA GCT TGG AAA TTG TCA CAG TGG GAT TTG GTG GAA Tyr Arg Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu 1610 1615 1620	5201
AAC TAT TTG GCA GCA GAT GGA AAA TCT ACA ACA TGG AGT GTC AGA CTG Asn Tyr Leu Ala Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg Leu 1625 1630 1635	5249
GGA CAG CTA TTA TTA TCA GCC AAA AAA AGA GAT ATC ACA GCT TTT TAT Gly Gln Leu Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr 1640 1645 1650 1655	5297
GAC TCA CTG AAA CTA GTG AGA GCA GAA CAA ATT GTA CCT CTT TCA GCT Asp Ser Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala 1660 1665 1670	5345
GCA AGC TTT GAA AGA GGC TCC TAC CAA CGA GGA TAT GAA TAT ATT GTG Ala Ser Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val 1675 1680 1685	5393
AGA TTG CAC ATG TTA TGT GAG TTG GAG CAT AGC ATC AAA CCA CTT TTC Arg Leu His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe 1690 1695 1700	5441
CAG CAT TCT CCA GGT GAC AGT TCT CAA GAA GAT TCT CTA AAC TGG GTA Gln His Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val 1705 1710 1715	5489
GCT CGA CTA GAA ATG ACC CAG AAT TCC TAC AGA GCC AAG GAG CCT ATC Ala Arg Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile 1720 1725 1730 1735	5537
CTG GCT CTC CGG AGG GCT TTA CTA AGC CTC AAC AAA AGA CCA GAT TAC Leu Ala Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr 1740 1745 1750	5585
AAT GAA ATG GTT GGA GAA TGC TGG CTG CAG AGT GCC AGG GTA GCT AGA Asn Glu Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg 1755 1760 1765	5633
AAG GCT GGT CAC CAC CAG ACA GCC TAC AAT GCT CTC CTT AAT GCA GGG Lys Ala Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly 1770 1775 1780	5681
GAA TCA CGA CTC GCT GAA CTG TAC GTG GAA AGG GCA AAG TGG CTC TGG Glu Ser Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp 1785 1790 1795	5729
TCC AAG GGT GAT GTT CAC CAG GCA CTA ATT GTT CTT CAA AAA GGT GTT Ser Lys Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val 1800 1805 1810 1815	5777
GAA TTA TGT TTT CCT GAA AAT GAA ACC CCA CCT GAG GGT AAG AAC ATG Glu Leu Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met 1820 1825 1830	5825
TTA ATC CAT GGT CGA GCT ATG CTA CTA GTG GGC CGA TTT ATG GAA GAA Leu Ile His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu 1835 1840 1845	5873

-59-

ACA GCT AAC TTT GAA AGC AAT GCA ATT ATG AAA AAA TAT AAG GAT GTG Thr Ala Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val 1850 1855 1860	5921
ACC GCG TGC CTG CCA GAA TGG GAG GAT GGG CAT TTT TAC CTT GCC AAG Thr Ala Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys 1865 1870 1875	5969
TAC TAT GAC AAA TTG ATG CCC ATG GTC ACA GAC AAC AAA ATG GAA AAG Tyr Tyr Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys 1880 1885 1890 1895	6017
CAA GGT GAT CTC ATC CGG TAT ATA GTT CTT CAT TTT GGC AGA TCT CTA Gln Gly Asp Leu Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu 1900 1905 1910	6065
CAA TAT GGA AAT CAG TTC ATA TAT CAG TCA ATG CCA CGA ATG TTA ACT Gln Tyr Gly Asn Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr 1915 1920 1925	6113
CTA TGG CTT GAT TAT GGT ACA AAG GCA TAT GAA TGG GAA AAA GCT GGC Leu Trp Leu Asp Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala Gly 1930 1935 1940	6161
CGC TCC GAT CGT GTA CAA ATG AGG AAT GAT TTG GGT AAA ATA AAC AAG Arg Ser Asp Arg Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys 1945 1950 1955	6209
GTT ATC ACA GAG CAT ACA AAC TAT TTA GCT CCA TAT CAA TTT TTG ACT Val Ile Thr Glu His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr 1960 1965 1970 1975	6257
GCT TTT TCA CAA TTG ATC TCT CGA ATT TGT CAT TCT CAC GAT GAA GTT Ala Phe Ser Gln Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val 1980 1985 1990	6305
TTT GTT GTG CTT GAT GGA AAT AAT AGC CAA GTA TTT CTA GCC TAT CCT Phe Val Val Leu Asp Gly Asn Asn Ser Gln Val Phe Leu Ala Tyr Pro 1995 2000 2005	6353
CAA CAA GCA ATG TGG ATG ATG ACA GCT GTG TCA AAG TCA TCT TAT CCC Gln Gln Ala Met Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr Pro 2010 2015 2020	6401
ATG CGT GTG AAC AGA TGC AAG GAA ATC CTC AAT AAA GCT ATT CAT ATG Met Arg Val Asn Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His Met 2025 2030 2035	6449
AAA AAA TCC TTA GAG AAG TTT GTT GGA GAT GCA ACT CGC CTA ACA GAT Lys Lys Ser Leu Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr Asp 2040 2045 2050 2055	6497
AAG CTT CTA GAA TTG TGC AAT AAA CCG GTG GAA ATT CTT GCT TCT CTT Lys Leu Leu Glu Leu Cys Asn Lys Pro Val Glu Ile Leu Ala Ser Leu 2060 2065 2070	6545
CAG AAA CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC TAC Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr 2075 2080 2085	6593
ATC ATG ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA CTA Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu 2090 2095 2100	6641
ATG GAA TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu 2105 2110 2115	6689

-60-

TCT CGT AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu 2120 2125 2130 2135	6737
AAT GAT GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu 2140 2145 2150	6785
AGA CCT ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr 2155 2160 2165	6833
GGA AAA GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser 2170 2175 2180	6881
GAA AAA CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro 2185 2190 2195	6929
ATT TTT CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp 2200 2205 2210 2215	6977
TAC AGT AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met 2220 2225 2230	7025
GTT GGT TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu 2235 2240 2245	7073
TTT GAT TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu 2250 2255 2260	7121
TTC AAT AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg 2265 2270 2275	7169
CTG ACT CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly 2280 2285 2290 2295	7217
CTT TTT CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT CAG Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln 2300 2305 2310	7265
CGA GAG CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT CTT Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu 2315 2320 2325	7313
GTG GAA TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG AAT Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn 2330 2335 2340	7361
GAA ACT GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT GAC Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp 2345 2350 2355	7409
ATT GAG CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG ACA Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr 2360 2365 2370 2375	7457
GGA CTG CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA GAA Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu 2380 2385 2390	7505

-61-

GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT CCA 7553
 Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro
 2395 2400 2405

TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA AGTAAAAAAA 7609
 Tyr Met

AAAAAAAAAA AA 7621

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2409 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser
 1 5 10 15
 Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu
 20 25 30
 Met Ser Met Gln Lys Leu Glu Phe Ile Glu Val Thr Leu Leu Thr Val
 35 40 45
 Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu
 50 55 60
 Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile
 65 70 75 80
 Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly
 85 90 95
 Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu
 100 105 110
 Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu
 115 120 125
 Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu
 130 135 140
 Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys
 145 150 155 160
 Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp
 165 170 175
 Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe
 180 185 190
 Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val
 195 200 205
 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp
 210 215 220
 Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser
 225 230 235 240

-62-

Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu
 245 250 255
 Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350
 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Thr Val Leu Lys Ser Cys Arg
 370 375 380
 Arg Leu Leu Glu Ser Val Gln Lys Arg Thr Gly Gly Asn Ile Asp Lys
 385 390 395 400
 Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser Ser
 405 410 415
 Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu Pro
 420 425 430
 Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr Phe
 435 440 445
 Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr Ser
 450 455 460
 Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro Arg
 465 470 475 480
 Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu Gln
 485 490 495
 Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe Ile
 500 505 510
 Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu Ile
 515 520 525
 Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala Ser
 530 535 540
 Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu Thr
 545 550 555 560
 Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu Phe
 565 570 575
 Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser Gln
 580 585 590

-63-

Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys Lys
 595 600 605

Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His Leu
 610 615 620

Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala Val
 625 630 635 640

Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val Arg
 645 650 655

Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp Ser
 660 665 670

Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu Ala
 675 680 685

Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr Leu
 690 695 700

Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu Val
 705 710 715 720

Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser Ala
 725 730 735

Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala Ala
 740 745 750

Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro Ile
 755 760 765

Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala Leu
 770 775 780

Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val Ala
 785 790 795 800

His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn Val
 805 810 815

Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln Val
 820 825 830

Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala Leu
 835 840 845

Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile Leu
 850 855 860

Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys Ser
 865 870 875 880

Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr Glu
 885 890 895

Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His Asn
 900 905 910

Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn Gly
 915 920 925

Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln Gly
 930 935 940

-64-

Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln Pro
 945 950 955 960
 Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser Ser
 965 970 975
 Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met Ser
 980 985 990
 Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val Lys
 995 1000 1005
 Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe Pro
 1010 1015 1020
 Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp His
 1025 1030 1035 1040
 Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu Pro
 1045 1050 1055
 Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr Leu
 1060 1065 1070
 Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile Tyr
 1075 1080 1085
 Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu Gln
 1090 1095 1100
 Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr Leu
 1105 1110 1115 1120
 Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg Ile
 1125 1130 1135
 His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu Lys
 1140 1145 1150
 Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile Ser
 1155 1160 1165
 Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser Gln
 1170 1175 1180
 Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile Asp
 1185 1190 1195 1200
 Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp Phe
 1205 1210 1215
 Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu Leu
 1220 1225 1230
 Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala
 1235 1240 1245
 Pro Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp
 1250 1255 1260
 Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg Arg
 1265 1270 1275 1280
 Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr Arg
 1285 1290 1295

-65-

Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys Pro
 1300 1305 1310
 Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser
 1315 1320 1325
 Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser Lys
 1330 1335 1340
 Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val Thr
 1345 1350 1355 1360
 Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys Asn
 1365 1370 1375
 Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu Lys
 1380 1385 1390
 His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu
 1395 1400 1405
 Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Thr
 1410 1415 1420
 Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro
 1425 1430 1435 1440
 His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr Val
 1445 1450 1455
 Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile Pro
 1460 1465 1470
 Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg
 1475 1480 1485
 Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile
 1490 1495 1500
 Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His Glu
 1505 1510 1515 1520
 Pro Asp Gly Val Ser Gly Val Ser Ala Ile Arg Lys Ala Glu Pro Ser
 1525 1530 1535
 Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg Asp
 1540 1545 1550
 Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile
 1555 1560 1565
 Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln Leu
 1570 1575 1580
 Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser Glu
 1585 1590 1595 1600
 Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu
 1605 1610 1615
 Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser
 1620 1625 1630
 Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys
 1635 1640 1645

-66-

Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu
 1650 1655 1660

Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln
 1665 1670 1675 1680

Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu Glu
 1685 1690 1695

His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln
 1700 1705 1710

Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser
 1715 1720 1725

Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser
 1730 1735 1740

Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu
 1745 1750 1755 1760

Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr
 1765 1770 1775

Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val
 1780 1785 1790

Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu
 1795 1800 1805

Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu Thr
 1810 1815 1820

Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu
 1825 1830 1835 1840

Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile
 1845 1850 1855

Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu Asp
 1860 1865 1870

Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met Val
 1875 1880 1885

Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val
 1890 1895 1900

Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln
 1905 1910 1915 1920

Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala
 1925 1930 1935

Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg Asn
 1940 1945 1950

Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr Leu
 1955 1960 1965

Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile
 1970 1975 1980

Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn Ser
 1985 1990 1995 2000

-67-

Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr Ala
 2005 2010 2015
 Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu Ile
 2020 2025 2030
 Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val Gly
 2035 2040 2045
 Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro
 2050 2055 2060
 Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys
 2065 2070 2075 2080
 Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp
 2085 2090 2095
 Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys
 2100 2105 2110
 Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile Arg
 2115 2120 2125
 Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp
 2130 2135 2140
 Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys
 2145 2150 2155 2160
 Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu
 2165 2170 2175
 Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe
 2180 2185 2190
 Leu Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr
 2195 2200 2205
 Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg
 2210 2215 2220
 Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp
 2225 2230 2235 2240
 Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val
 2245 2250 2255
 His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val
 2260 2265 2270
 Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly Met
 2275 2280 2285
 Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr
 2290 2295 2300
 Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys
 2305 2310 2315 2320
 Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly
 2325 2330 2335
 His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys
 2340 2345 2350

-68-

Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile
 2355 2360 2365

Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His
 2370 2375 2380

Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln
 2385 2390 2395 2400

Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2405

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2835 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:

- (B) CLONE: 517

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2610

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTG GAA GCA GCT TGG AAA TTG TCA CAG TGG GAT TTG GTG GAA AAC TAT	48
Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr	
1 5 10 15	
TTG GCA GCA GAT GGA AAA TCT ACA ACA TGG AGT GTC AGA CTG GGA CAG	96
Leu Ala Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg Leu Gly Gln	
20 25 30	
CTA TTA TTA TCA GCC AAA AAA AGA GAT ATC ACA GCT TTT TAT GAC TCA	144
Leu Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser	
35 40 45	
CTG AAA CTA GTG AGA GCA GAA CAA ATT GTA CCT CTT TCA GCT GCA AGC	192
Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala Ala Ser	
50 55 60	
TTT GAA AGA GGC TCC TAC CAA CGA GGA TAT GAA TAT ATT GTG AGA TTG	240
Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu	
65 70 75 80	
CAC ATG TTA TGT GAG TTG GAG CAT AGC ATC AAA CCA CTT TTC CAG CAT	288
His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe Gln His	
85 90 95	
TCT CCA GGT GAC AGT TCT CAA GAA GAT TCT CTA AAC TGG GTA GCT CGA	336
Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg	
100 105 110	
CTA GAA ATG ACC CAG AAT TCC TAC AGA GCC AAG GAG CCT ATC CTG GCT	384
Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala	
115 120 125	

-69-

CTC CGG AGG GCT TTA CTA AGC CTC AAC AAA AGA CCA GAT TAC AAT GAA Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu 130 135 140	432
ATG GTT GGA GAA TGC TGG CTG CAG AGT GCC AGG GTA GCT AGA AAG GCT Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala 145 150 155 160	480
GGT CAC CAC CAG ACA GCC TAC AAT GCT CTC CTT AAT GCA GGG GAA TCA Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser 165 170 175	528
CGA CTC GCT GAA CTG TAC GTG GAA AGG GCA AAG TGG CTC TGG TCC AAG Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys 180 185 190	576
GGT GAT GTT CAC CAG GCA CTA ATT GTT CTT CAA AAA GGT GTT GAA TTA Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu 195 200 205	624
TGT TTT CCT GAA AAT GAA ACC CCA CCT GAG GGT AAG AAC ATG TTA ATC Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile 210 215 220	672
CAT GGT CGA GCT ATG CTA CTA GTG GGC CGA TTT ATG GAA GAA ACA GCT His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala 225 230 235 240	720
AAC TTT GAA AGC AAT GCA ATT ATG AAA AAA TAT AAG GAT GTG ACC GCG Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala 245 250 255	768
TGC CTG CCA GAA TGG GAG GAT GGG CAT TTT TAC CTT GCC AAG TAC TAT Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr 260 265 270	816
GAC AAA TTG ATG CCC ATG GTC ACA GAC AAC AAA ATG GAA AAG CAA GGT Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly 275 280 285	864
GAT CTC ATC CGG TAT ATA GTT CTT CAT TTT GGC AGA TCT CTA CAA TAT Asp Leu Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu Gln Tyr 290 295 300	912
GGA AAT CAG TTC ATA TAT CAG TCA ATG CCA CGA ATG TTA ACT CTA TGG Gly Asn Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr Leu Trp 305 310 315 320	960
CTT GAT TAT GGT ACA AAG TCA TAT GAA TGG GAA AAA GCT GGC CGC TCC Leu Asp Tyr Gly Thr Lys Ser Tyr Glu Trp Glu Lys Ala Gly Arg Ser 325 330 335	1008
GAT CGT GTA CAA ATG AGG AAT GAT TTG GGT AAA ATA AAC AAG GTT ATC Asp Arg Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys Val Ile 340 345 350	1056
ACA GAG CAT ACA AAC TAT TTA GCT CCA TAT CAA TTT TTG ACT GCT TTT Thr Glu His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe 355 360 365	1104
TCA CAA TTG ATC TCT CGA ATT TGT CAT TCT CAC GAT GAA GTT TTT GTT Ser Gln Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val Phe Val 370 375 380	1152

-70-

GTC Val 385	TTG Leu	ATG Met	GAA Glu	ATA Ile	ATA Ile	GCC Ala	AAA Lys	GTA Val	TTT Phe	CTA Leu	GCC Ala	TAT Tyr	CCT Pro	CAA Gln	CAA Gln	1200
GCA Ala	ATG Met	TGG Trp	ATG Met	ATG Met	ACA Thr	GCT Ala	GTG Val	TCA Ser	AAG Lys	TCA Ser	TCT Ser	TAT Tyr	CCC Pro	ATG Met	CGT Arg	1248
GTG Val	AAC Asn	AGA Arg	TGC Cys	AAG Lys	GAA Glu	ATC Ile	CTC Leu	AAT Asn	AAA Lys	GCT Ala	ATT Ile	CAT His	ATG Met	AAA Lys	AAA Lys	1296
TCC Ser	TTA Leu	GAG Glu	AAG Lys	TTT Phe	GTT Val	GGA Gly	GAT Asp	GCA Ala	ACT Thr	CGC Arg	CTA Leu	ACA Thr	GAT Asp	AAG Lys	CTT Leu	1344
CTA Leu	GAA Glu	TTG Leu	TGC Cys	AAT Asn	AAA Lys	CCG Pro	GTT Val	GAT Asp	GGA Gly	AGT Ser	AGT Ser	TCC Ser	ACA Thr	TTA Leu	AGC Ser	1392
ATG Met	AGC Ser	ACT Thr	CAT His	TTT Phe	AAA Lys	ATG Met	CTT Leu	AAA Lys	AAG Lys	CTG Leu	GTA Val	GAA Glu	GAA Glu	GCA Ala	ACA Thr	1440
TTT Phe	AGT Ser	GAA Glu	ATC Ile	CTC Leu	ATT Ile	CCT Pro	CTA Leu	CAA Gln	TCA Ser	GTC Val	ATG Met	ATA Ile	CCT Pro	ACA Thr	CTT Leu	1488
CCA Pro	TCA Ser	ATT Ile	CTG Leu	GGT Gly	ACC Thr	CAT His	GCT Ala	AAC Asn	CAT His	GCT Ala	AGC Ser	CAT His	GAA Glu	CCA Pro	TTT Phe	1536
CCT Pro	GGA Gly	CAT His	TGG Trp	GCC Ala	TAT Tyr	ATT Ile	GCA Ala	GGG Gly	TTT Phe	GAT Asp	GAT Asp	ATG Met	GTG Val	GAA Glu	ATT Ile	1584
CTT Leu	GCT Ala	TCT Ser	CTT Leu	CAG Gln	AAA Lys	CCA Pro	AAG Lys	AAG Lys	ATT Ile	TCT Ser	TTA Leu	AAA Lys	GGC Gly	TCA Ser	GAT Asp	1632
GGA Gly	AAG Lys	TTC Phe	TAC Tyr	ATC Ile	ATG Met	ATG Met	TGT Cys	AAG Lys	CCA Pro	AAA Lys	GAT Asp	GAC Asp	CTG Leu	AGA Arg	AAG Lys	1680
GAT Asp	TGT Cys	AGA Arg	CTA Leu	ATG Met	GAA Glu	TTC Phe	AAT Asn	TCC Ser	TTG Leu	ATT Ile	AAT Asn	AAG Lys	TGC Cys	TTA Leu	AGA Arg	1728
AAA Lys	GAT Asp	GCA Ala	GAG Glu	TCT Ser	CGT Arg	AGA Arg	AGA Arg	GAA Glu	CTT Leu	CAT His	ATT Ile	CGA Arg	ACA Thr	TAT Tyr	GCA Ala	1776
GTT Val	ATT Ile	CCA Pro	CTA Leu	AAT Asn	GAT Asp	GAA Glu	TGT Cys	GGG Gly	ATT Ile	ATT Ile	GAA Glu	TGG Trp	GTG Val	AAC Asn	AAC Asn	1824
ACT Thr	GCT Ala	GGT Gly	TTG Leu	AGA Arg	CCT Pro	ATT Ile	CTG Leu	ACC Thr	AAA Lys	CTA Leu	TAT Tyr	AAA Lys	GAA Glu	AAG Lys	GGA Gly	1872
GTG Val	TAT Tyr	ATG Met	ACA Thr	GGA Gly	AAA Lys	GAA Glu	CTT Leu	CGC Arg	CAG Gln	TGT Cys	ATG Met	CTA Leu	CCA Pro	AAG Lys	TCA Ser	1920
GCA Ala	GCT Ala	TTA Leu	TCT Ser	GAA Glu	AAA Lys	CTC Leu	AAA Lys	GTA Val	TTC Phe	CGA Arg	GAA Glu	TTT Phe	CTC Leu	CTG Leu	CCC Pro	1968

-71-

AGG CAT CCT CCT ATT TTT CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp 660 665 670	2016
CCT ACA TCA TGG TAC AGT AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala 675 680 685	2064
GTA ATG TCA ATG GTT GGT TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly 690 695 700	2112
GAA AAT ATT CTC TTT GAT TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp 705 710 715 720	2160
TTC AAT TGT CTT TTC AAT AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile 725 730 735	2208
GTG CCA TTT CGC CTG ACT CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met 740 745 750	2256
GGA ACA GAG GGT CTT TTT CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu 755 760 765	2304
ATG CGT GAT CAG CGA GAG CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu 770 775 780	2352
CAT GAT CCT CTT GTG GAA TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys 785 790 795 800	2400
GCG CCA CTG AAT GAA ACT GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr 805 810 815	2448
CAT GTT CTT GAC ATT GAG CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg 820 825 830	2496
AAT AGA GTG ACA GGA CTG CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr 835 840 845	2544
CTT ATA CAA GAA GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu 850 855 860	2592
GGT TGG ACT CCA TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA Gly Trp Thr Pro Tyr Met 865 870	2640
ATAATCTAAA AGTAATGCAT TTGGTATGAA TCTGTGGTTG TATCTGTTCA ATTCTAAAGT	2700
ACAACATAAA TTTACGTTCT CAGCAACTGT TATTTCTCTC TGATCATTAA TTATATGTAA	2760
AATAATATAC ATTCAATTAT TAAGAAATAA ACTGCTTTCT TAATAAAAAA AAAAAAAAAA	2820
AAAAAAAAA AAAAA	2835

-72-

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 870 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr
 1           5           10           15
Leu Ala Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg Leu Gly Gln
 20           25           30
Leu Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser
 35           40           45
Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala Ala Ser
 50           55           60
Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu
 65           70           75           80
His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe Gln His
 85           90           95
Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg
 100          105          110
Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala
 115          120          125
Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu
 130          135          140
Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala
 145          150          155          160
Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser
 165          170          175
Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys
 180          185          190
Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu
 195          200          205
Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile
 210          215          220
His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala
 225          230          235          240
Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala
 245          250          255
Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr
 260          265          270
Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly
 275          280          285

```

-73-

Asp Leu Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu Gln Tyr
 290 295 300
 Gly Asn Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr Leu Trp
 305 310 315 320
 Leu Asp Tyr Gly Thr Lys Ser Tyr Glu Trp Glu Lys Ala Gly Arg Ser
 325 330 335
 Asp Arg Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys Val Ile
 340 345 350
 Thr Glu His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe
 355 360 365
 Ser Gln Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val Phe Val
 370 375 380
 Val Leu Met Glu Ile Ile Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln
 385 390 395 400
 Ala Met Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr Pro Met Arg
 405 410 415
 Val Asn Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His Met Lys Lys
 420 425 430
 Ser Leu Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu
 435 440 445
 Leu Glu Leu Cys Asn Lys Pro Val Asp Gly Ser Ser Ser Thr Leu Ser
 450 455 460
 Met Ser Thr His Phe Lys Met Leu Lys Lys Leu Val Glu Glu Ala Thr
 465 470 475 480
 Phe Ser Glu Ile Leu Ile Pro Leu Gln Ser Val Met Ile Pro Thr Leu
 485 490 495
 Pro Ser Ile Leu Gly Thr His Ala Asn His Ala Ser His Glu Pro Phe
 500 505 510
 Pro Gly His Trp Ala Tyr Ile Ala Gly Phe Asp Asp Met Val Glu Ile
 515 520 525
 Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp
 530 535 540
 Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys
 545 550 555 560
 Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg
 565 570 575
 Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala
 580 585 590
 Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn
 595 600 605
 Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly
 610 615 620
 Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser
 625 630 635 640

-74-

Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro
 645 650 655
 Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp
 660 665 670
 Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala
 675 680 685
 Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly
 690 695 700

 Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp
 705 710 715 720
 Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile
 725 730 735
 Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met
 740 745 750
 Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu
 755 760 765
 Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu
 770 775 780
 His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys
 785 790 795 800
 Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr
 805 810 815
 His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg
 820 825 830
 Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr
 835 840 845
 Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu
 850 855 860
 Gly Trp Thr Pro Tyr Met
 865 870

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
- (B) CLONE: Primer oDH15a

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: group(15, 18, 24, 30)
- (D) OTHER INFORMATION: The nucleotides at these positions are inosines.

-75-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCAGACGGAT CCGGNWCNGA YGGNAAYHTN TAY

33

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: group(15, 18, 24)
- (D) OTHER INFORMATION: The nucleotides at these positions are inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCAGACGGAT CCGGNWCNGA YGGNAAY

27

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

- (B) CLONE: Primer oDH16

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 24
- (D) OTHER INFORMATION: The nucleotides at these positions are inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCAGACGAAT TCRCARTYRA ARTCNACRTG

30

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 41 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

- (B) CLONE: Primer oDH17a

-76-

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: group(21, 24, 27, 30)
 (D) OTHER INFORMATION: The nucleotides at these positions are
inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
GCAGACGGAT CCAARTTYCC NCCNRTNYTN TAYSARTGGT T 41

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:
 (B) CLONE: Primer oDH17b

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: group(24, 27, 30, 33)
 (D) OTHER INFORMATION: The nucleotides at these positions are
inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
GCAGACGAAT CCAACCAYTS RTANARNAYN GGNGGAAAYT T 41

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:
 (B) CLONE: Primer oDH18a

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: group(15, 18, 21, 24, 30)
 (D) OTHER INFORMATION: The nucleotides at these positions are
inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
GCAGACGGAT CCYTNGGNYT NGGNGAYCGN CA 32

-77-

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oDH18b

- (ix) FEATURE:
 - (A) NAME/KEY: modified_base
 - (B) LOCATION: group(15, 18, 21)
 - (D) OTHER INFORMATION: The nucleotides at these positions are inosines.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCAGACGGAT CCYTNGGNYT NGGNGAYAGR CA

32

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oDH23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GACGCAGAAT TCACCACTCA AAGAATCAAA GAG

33

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer mo3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTACAGAGCC AAGGAG

16

-78-

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer mo6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCGAGCTATG CTACTAGTGG GC

22

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oHT9-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCAGTAAACT TGCTTTC

17

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oHT9-4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTTGCGGCC TTCCAATATC

20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7440 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

-79-

(vii) IMMEDIATE SOURCE:
(B) CLONE: MCCS1beta

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..7437

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATG GGT CAT GCT GTG GAA TGG CCA GTG GTC ATG AGC CGA TTT TTA AGT	48
Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser	
1 5 10 15	
CAA TTA GAT GAA CAC ATG GGA TAT TTA CAA TCA GCT CCT TTG CAG TTG	96
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu	
20 25 30	
ATG AGT ATG CAA AAA TTA GAA TTT ATT GAA GTC ACT TTA TTA ACG GTT	144
Met Ser Met Gln Lys Leu Glu Phe Ile Glu Val Thr Leu Leu Thr Val	
35 40 45	
CTT ACT CGT ATT ATT GCA ATT GTG TTT TTT AGA AGG CAA GAA CTC TTA	192
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu	
50 55 60	
CTT TGG CAG ATA GGT TGT GTT CTG CTA GAG TAT GGT AGT CCA AAA ATT	240
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile	
65 70 75 80	
AAA TCC CTA GCA ATT AGC TTT TTA ACA GAA CTT TTT CAG CTT GGA GGA	288
Lys Ser Leu Ala Ile Ser Phe Leu Thr Leu Phe Gln Leu Gly Gly	
85 90 95	
CTA CCA GCA CAA CCA GCT AGC ACT TTT TTC AGC TCA TTT TTG GAA TTA	336
Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu	
100 105 110	
TTA AAA CAC CTT GTA GAA ATG GAT ACT GAC CAA TTG AAA CTC TAT GAA	384
Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu	
115 120 125	
GAG CCA TTA TCA AAG CTG ATA AAG ACA CTA TTT CCC TTT GAA GCA GAA	432
Glu Pro Leu Ser Lys Leu Lys Lys Thr Leu Phe Pro Phe Glu Ala Glu	
130 135 140	
GCT TAT AGA AAT ATT GAA CCT GTC TAT TTA AAT ATG CTG CTG GAA AAA	480
Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys	
145 150 155 160	
CTC TGT GTC ATG TTT GAA GAC GGT GTG CTC ATG CGG CTT AAG TCT GAT	528
Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp	
165 170 175	
TTG CTA AAA GCA GCT TTG TGC CAT TTA CTG CAG TAT TTC CTT AAA TTT	576
Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe	
180 185 190	
GTG CCA GCT GGG TAT GAA TCT GCT TTA CAA GTC AGG AAG GTC TAT GTG	624
Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val	
195 200 205	
AGA AAT ATT TGT AAA GCT CTT TTG GAT GTG CTT GGA ATT GAG GTA GAT	672
Arg Asn Ile Cys Lys Ala Leu Asp Val Leu Gly Ile Glu Val Asp	
210 215 220	

-80-

GCA Ala 225	GAG Glu 225	TAC Tyr 225	TTG Leu 225	TTG Leu 225	GGC Gly 230	CCA Pro 230	CTT Leu 230	TAT Tyr 230	GCA Ala 235	GCT Ala 235	TTG Leu 235	AAA Lys 235	ATG Met 240	GAA Glu 240	AGT Ser 240	720
ATG Met 245	GAA Glu 245	ATC Ile 245	ATT Ile 245	GAG Glu 245	GAG Glu 245	ATT Ile 245	CAA Gln 250	TGC Cys 250	CAA Gln 250	ACT Thr 250	CAA Gln 255	CAG Gln 255	GAA Glu 255	AAC Asn 255	CTC Leu 255	768
AGC Ser 260	AGT Ser 260	AAT Asn 260	AGT Ser 260	GAT Asp 260	GGA Gly 265	ATA Ile 265	TCA Ser 265	CCC Pro 265	AAA Lys 265	AGG Arg 270	CGT Arg 270	CGT Arg 270	CTC Leu 270	AGC Ser 270	TCG Ser 270	816
TCT Ser 275	CTA Leu 275	AAC Asn 275	CCT Pro 275	TCT Ser 280	AAA Lys 280	AGA Arg 280	GCA Ala 280	CCA Pro 280	AAA Lys 285	CAG Gln 285	ACT Thr 285	GAG Glu 285	GAA Glu 285	ATT Ile 285	AAA Lys 285	864
CAT His 290	GTG Val 290	GAC Asp 290	ATG Met 295	AAC Asn 295	CAA Gln 295	AAG Lys 295	AGC Ser 295	ATA Ile 300	TTA Leu 300	TGG Trp 300	AGT Ser 300	GCA Ala 300	CTG Leu 300	AAA Lys 300	CAG Gln 300	912
AAA Lys 305	GCT Ala 305	GAA Glu 305	TCC Ser 310	CTT Leu 310	CAG Gln 310	ATT Ile 310	TCC Ser 315	CTT Leu 315	GAA Glu 315	TAC Tyr 315	AGT Ser 315	GGC Gly 315	CTA Leu 315	AAG Lys 315	AAT Asn 320	960
CCT Pro 325	GTT Val 325	ATT Ile 325	GAG Glu 325	ATG Met 325	TTA Leu 325	GAA Glu 325	GGA Gly 330	ATT Ile 330	GCT Ala 330	GTT Val 330	GTC Val 330	TTA Leu 330	CAA Gln 330	CTG Leu 330	ACT Thr 330	1008
GCT Ala 340	CTG Leu 340	TGT Cys 340	ACT Thr 340	GTT Val 345	CAT His 345	TGT Cys 345	TCT Ser 345	CAT His 345	CAA Gln 350	AAC Asn 350	ATG Met 350	AAC Asn 350	TGC Cys 350	CGT Arg 350	ACT Thr 350	1056
TTC Phe 355	AAG Lys 355	GAC Asp 355	TGT Cys 360	CAA Gln 360	CAT His 360	AAA Lys 360	TCC Ser 360	AAG Lys 365	AAG Lys 365	AAA Lys 365	CCT Pro 365	TCT Ser 365	GTA Val 365	GTG Val 365	ATA Ile 365	1104
ACT Thr 370	TGG Trp 370	ATG Met 375	TCA Ser 375	TTG Leu 375	GAT Asp 375	TTT Phe 375	TAC Tyr 375	ACA Thr 380	ACA Thr 380	GTG Val 380	CTT Leu 380	AAG Lys 380	AGC Ser 380	TGT Cys 380	AGA Arg 380	1152
AGG Arg 385	TTG Leu 385	TTA Leu 390	GAA Glu 390	TCT Ser 390	GTT Val 390	CAG Gln 390	AAA Lys 395	CGG Arg 395	ACT Thr 395	GGA Gly 395	GGC Gly 395	AAC Asn 400	ATT Ile 400	GAT Asp 400	AAG Lys 400	1200
GTG Val 405	GTG Val 405	AAA Lys 405	ATT Ile 405	TAT Tyr 405	GAT Asp 410	GCT Ala 410	TTG Leu 410	ATT Ile 410	TAT Tyr 410	ATG Met 415	CAA Gln 415	GTA Val 415	AAC Asn 415	AGT Ser 415	TCA Ser 415	1248
TTT Phe 420	GAA Glu 420	GAT Asp 420	CAT His 420	ATC Ile 425	CTG Leu 425	GAA Glu 425	GAT Asp 425	TTA Leu 425	TGT Cys 425	GGA Gly 430	ATG Met 430	CTC Leu 430	TCA Ser 430	CTT Leu 430	CCA Pro 430	1296
TGG Trp 435	ATT Ile 435	TAT Tyr 435	TCC Ser 440	CAT His 440	TCT Ser 440	GAT Asp 440	GAT Asp 440	GGC Gly 445	TGT Cys 445	TTA Leu 445	AAG Lys 445	TTG Leu 445	ACC Thr 445	ACA Thr 445	TTT Phe 445	1344
GCC Ala 450	GCT Ala 450	AAT Asn 450	CTT Leu 455	CTA Leu 455	ACA Thr 455	TTA Leu 455	AGC Ser 460	TGT Cys 460	AGG Arg 460	ATT Ile 460	TCA Ser 460	GAT Asp 460	AGC Ser 460	TAT Tyr 460	TCA Ser 460	1392
CCA Pro 465	CAG Gln 465	GCA Ala 470	CAA Gln 470	TCA Ser 470	CGA Arg 470	TGT Cys 470	GTG Val 475	TTT Phe 475	CTT Leu 475	CTG Leu 475	ACT Thr 475	CTG Leu 475	TTT Phe 475	CCA Pro 480	AGA Arg 480	1440
AGA Arg 485	ATA Ile 485	TTC Phe 485	CTT Leu 485	GAG Glu 485	TGG Trp 485	AGA Arg 490	ACA Thr 490	GCA Ala 490	GTT Val 490	TAC Tyr 490	AAC Asn 490	TGG Trp 490	GCC Ala 495	CTG Leu 495	CAG Gln 495	1488

-81-

AGC TCC CAT GAA GTA ATC CGG GCT AGT TGT GTT AGT GGA TTT TTT ATC Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe Ile 500 505 510	1536
TTA TTG CAG CAG CAG AAT TCT TGT AAC AGA GTT CCC AAG ATT CTT ATA Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu Ile 515 520 525	1584
GAT AAA GTC AAA GAT GAT TCT GAC ATT GTC AAG AAA GAA TTT GCT TCT Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala Ser 530 535 540	1632
ATA CTT GGT CAA CTT GTC TGT ACT CTT CAC GGC ATG TTT TAT CTG ACA Ile Leu Gly Gln Leu Cys Thr Leu His Gly Met Phe Tyr Leu Thr 545 550 555 560	1680
AGT TCT TTA ACA GAA CCT TTC TCT GAA CAC GGA CAT GTG GAC CTC TTC Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu Phe 565 570 575	1728
TGT AGG AAC TTG AAA GCC ACT TCT CAA CAT GAA TGT TCA TCT TCT CAA Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser Gln 580 585 590	1776
CTA AAA GCT TCT GTC TGC AAG CCA TTC CTT TTC CTA CTG AAA AAA AAA Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys Lys 595 600 605	1824
ATA CCT AGT CCA GTA AAA CTT GCT TTC ATA GAT AAT CTA CAT CAT CTT Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His Leu 610 615 620	1872
TGT AAG CAT CTT GAT TTT AGA GAA GAT GAA ACA GAT GTA AAA GCA GTT Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala Val 625 630 635 640	1920
CTT GGA ACT TTA TTA AAT TTA ATG GAA GAT CCA GAC AAA GAT GTT AGA Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val Arg 645 650 655	1968
GTG GCT TTT AGT GGA AAT ATC AAG CAC ATA TTG GAA TCC TTG GAC TCT Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp Ser 660 665 670	2016
GAA GAT GGA TTT ATA AAG GAG CTT TTT GTC TTA AGA ATG AAG GAA GCA Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu Ala 675 680 685	2064
TAT ACA CAT GCC CAA ATA TCA AGA AAT AAT GAG CTG AAG GAT ACC TTG Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr Leu 690 695 700	2112
ATT CTT ACA ACA GGG GAT ATT GGA AGG GCC GCA AAA GGA GAT TTG GTA Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu Val 705 710 715 720	2160
CCA TTT GCA CTC TTA CAC TTA TTG CAT TGT TTG TTA TCC AAG TCA GCA Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser Ala 725 730 735	2208
TCT GTC TCT GGA GCA GCA TAC ACA GAA ATT AGA GCT CTG GTT GCA GCT Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala Ala 740 745 750	2256
AAA AGT GTT AAA CTG CAA AGT TTT TTC AGC CAG TAT AAG AAA CCC ATC Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro Ile 755 760 765	2304

-82-

TGT CAG TTT TTG GTA GAA TCC CTT CAC TCT AGT CAG ATG ACA GCA CTT Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala Leu 770 775 780	2352
CCG AAT ACT CCA TGC CAG AAT GCT GAC GTG CGA AAA CAA GAT GTG GCT Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val Ala 785 790 795 800	2400
CAC CAG AGA GAA ATG GCT TTA AAT ACG TTG TCT GAA ATT GCC AAC GTT His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn Val 805 810 815	2448
TTC GAC TTT CCT GAT CTT AAT CGT TTT CTT ACT AGG ACA TTA CAA GTT Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln Val 820 825 830	2496
CTA CTA CCT GAT CTT GCT GCC AAA GCA AGC CCT GCA GCT TCT GCT CTC Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala Leu 835 840 845	2544
ATT CGA ACT TTA GGA AAA CAA TTA AAT GTC AAT CGT AGA GAG ATT TTA Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile Leu 850 855 860	2592
ATA AAC AAC TTC AAA TAT ATT TTT TCT CAT TTG GTC TGT TCT TGT TCC Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys Ser 865 870 875 880	2640
AAA GAT GAA TTA GAA CGT GCC CTT CAT TAT CTG AAG AAT GAA ACA GAA Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr Glu 885 890 895	2688
ATT GAA CTG GGG AGC CTG TTG AGA CAA GAT TTC CAA GGA TTG CAT AAT Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His Asn 900 905 910	2736
GAA TTA TTG CTG CGT ATT GGA GAA CAC TAT CAA CAG GTT TTT AAT GGT Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn Gly 915 920 925	2784
TTG TCA ATA CTT GCC TCA TTT GCA TCC AGT GAT GAT CCA TAT CAG GGC Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln Gly 930 935 940	2832
CCG AGA GAT ATC ATA TCA CCT GAA CTG ATG GCT GAT TAT TTA CAA CCC Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln Pro 945 950 955 960	2880
AAA TTG TTG GGC ATT TTG GCT TTT TTT AAC ATG CAG TTA CTG AGC TCT Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser Ser 965 970 975	2928
AGT GTT GGC ATT GAA GAT AAG AAA ATG GCC TTG AAC AGT TTG ATG TCT Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met Ser 980 985 990	2976
TTG ATG AAG TTA ATG GGA CCC AAA CAT GTC AGT TCT GTG AGG GTG AAG Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val Lys 995 1000 1005	3024
ATG ATG ACC ACA CTG AGA ACT GGC CTT CGA TTC AAG GAT GAT TTT CCT Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe Pro 1010 1015 1020	3072
GAA TTG TGT TGC AGA GCT TGG GAC TGC TTT GTT CGC TGC CTG GAT CAT Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp His 1025 1030 1035 1040	3120

-83-

GCT TGT CTG GGC TCC CTT CTC AGT CAT GTA ATA GTA GCT TTG TTA CCT Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu Pro 1045 1050 1055	3168
CTT ATA CAC ATC CAG CCT AAA GAA ACT GCA GCT ATC TTC CAC TAC CTC Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr Leu 1060 1065 1070	3216
ATA ATT GAA AAC AGG GAT GCT GTG CAA GAT TTT CTT CAT GAA ATA TAT Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile Tyr 1075 1080 1085	3264
TTT TTA CCT GAT CAT CCA GAA TTA AAA AAG ATA AAA GCC GTT CTC CAG Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu Gln 1090 1095 1100	3312
GAA TAC AGA AAG GAG ACC TCT GAG AGC ACT GAT CTT CAG ACA ACT CTT Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr Leu 1105 1110 1115 1120	3360
CAG CTC TCT ATG AAG GCC ATT CAA CAT GAA AAT GTC GAT GTT CGT ATT Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg Ile 1125 1130 1135	3408
CAT GCT CTT ACA AGC TTG AAG GAA ACC TTG TAT AAA AAT CAG GAA AAA His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu Lys 1140 1145 1150	3456
CTG ATA AAG TAT GCA ACA GAC AGT GAA ACA GTA GAA CCT ATT ATC TCA Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile Ser 1155 1160 1165	3504
CAG TTG GTG ACA GTG CTT TTG AAA GGT TGC CAA GAT GCA AAC TCT CAA Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser Gln 1170 1175 1180	3552
GCT CGG TTG CTC TGT GGG GAA TGT TTA GGG GAA TTG GGG GCG ATA GAT Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile Asp 1185 1190 1195 1200	3600
CCA GGT CGA TTA GAT TTC TCA ACA ACT GAA ACT CAA GGA AAA GAT TTT Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp Phe 1205 1210 1215	3648
ACA TTT GTG ACT GGA GTA GAA GAT TCA AGC TTT GCC TAT GGA TTA TTG Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu Leu 1220 1225 1230	3696
ATG GAG CTA ACA AGA GCT TAC CTT GCG TAT GCT GAT AAT AGC CGA GCT Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala 1235 1240 1245	3744
CCA GAT TCA GCT GCC TAT GCC ATT CAG GAG TTG CTT TCT ATT TAT GAC Pro Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp 1250 1255 1260	3792
TGT AGA GAG ATG GAG ACC AAC GGC CCA GGT CAC CAA TTG TGG AGG AGA Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg Arg 1265 1270 1275 1280	3840
TTT CCT GAG CAT GTT CGG GAA ATA CTA GAA CCT CAT CTA AAT ACC AGA Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr Arg 1285 1290 1295	3888
TAC AAG AGT TCT CAG AAG TCA ACC GAT TGG TCT GGA GTA AAG AAG CCA Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys Pro 1300 1305 1310	3936

-84-

ATT TAC TTA AGT AAA TTG GGT AGT AAC TTT GCA GAA TGG TCA GCA TCT Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser 1315 1320 1325	3984
TGG GCA GGT TAT CTT ATT ACA AAG GTT CGA CAT GAT CTT GCC AGT AAA Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser Lys 1330 1335 1340	4032
ATT TTC ACC TGC TGT AGC ATT ATG ATG AAG CAT GAT TTC AAA GTG ACC Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val Thr 1345 1350 1355 1360	4080
ATC TAT CTT CTT CCA CAT ATT CTG GTG TAT GTC TTA CTG GGT TGT AAT Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys Asn 1365 1370 1375	4128
CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT ATG GCA GTT CTA AAG Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu Lys 1380 1385 1390	4176
CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC ATT GCA TCT GAT CTG His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu 1395 1400 1405	4224
TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC ACA Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Thr 1410 1415 1420	4272
CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT CCA Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro 1425 1430 1435 1440	4320
CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT GTG His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr Val 1445 1450 1455	4368
GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA CCC Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile Pro 1460 1465 1470	4416
CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA CGA Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg 1475 1480 1485	4464
GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT ATT Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile 1490 1495 1500	4512
CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT GAA Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His Glu 1505 1510 1515 1520	4560
CCT GAT GGA GTG TCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA TCT Pro Asp Gly Val Ser Gly Val Ser Ala Ile Arg Lys Ala Glu Pro Ser 1525 1530 1535	4608
CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG GAT Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg Asp 1540 1545 1550	4656
GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG ATC Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile 1555 1560 1565	4704
ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG CTG Ile His Tyr His Gly Val Lys Ser Met Leu Gly Leu Gly Gln Leu 1570 1575 1580	4752

-85-

TCT ACT GTT ATC ACT CAG GTG AAT GGA GTG CAT GCT AAC AGG TCC GAG Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser Glu 1585 1590 1595 1600	4800
TGG ACA GAT GAA TTA AAC ACG TAC AGA GTG GAA GCA GCT TGG AAA TTG Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu 1605 1610 1615	4848
TCA CAG TGG GAT TTG GTG GAA AAC TAT TTG GCA GCA GAT GGA AAA TCT Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser 1620 1625 1630	4896
ACA ACA TGG AGT GTC AGA CTG GGA CAG CTA TTA TTA TCA GCC AAA AAA Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys 1635 1640 1645	4944
AGA GAT ATC ACA GCT TTT TAT GAC TCA CTG AAA CTA GTG AGA GCA GAA Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu 1650 1655 1660	4992
CAA ATT GTA CCT CTT TCA GCT GCA AGC TTT GAA AGA GGC TCC TAC CAA Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln 1665 1670 1675 1680	5040
CGA GGA TAT GAA TAT ATT GTG AGA TTG CAC ATG TTA TGT GAG TTG GAG Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu Glu 1685 1690 1695	5088
CAT AGC ATC AAA CCA CTT TTC CAG CAT TCT CCA GGT GAC AGT TCT CAA His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln 1700 1705 1710	5136
GAA GAT TCT CTA AAC TGG GTA GCT CGA CTA GAA ATG ACC CAG AAT TCC Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser 1715 1720 1725	5184
TAC AGA GCC AAG GAG CCT ATC CTG GCT CTC CGG AGG GCT TTA CTA AGC Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser 1730 1735 1740	5232
CTC AAC AAA AGA CCA GAT TAC AAT GAA ATG GTT GGA GAA TGC TGG CTG Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu 1745 1750 1755 1760	5280
CAG AGT GCC AGG GTA GCT AGA AAG GCT GGT CAC CAC CAG ACA GCC TAC Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr 1765 1770 1775	5328
AAT GCT CTC CTT AAT GCA GGG GAA TCA CGA CTC GCT GAA CTG TAC GTG Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val 1780 1785 1790	5376
GAA AGG GCA AAG TGG CTC TGG TCC AAG GGT GAT GTT CAC CAG GCA CTA Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu 1795 1800 1805	5424
ATT GTT CTT CAA AAA GGT GTT GAA TTA TGT TTT CCT GAA AAT GAA ACC Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu Thr 1810 1815 1820	5472
CCA CCT GAG GGT AAG AAC ATG TTA ATC CAT GGT CGA GCT ATG CTA CTA Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu 1825 1830 1835 1840	5520
GTG GGC CGA TTT ATG GAA GAA ACA GCT AAC TTT GAA AGC AAT GCA ATT Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile 1845 1850 1855	5568

-86-

ATG AAA AAA TAT AAG GAT GTG ACC GCG TGC CTG CCA GAA TGG GAG GAT Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu Asp 1860 1865 1870	5616
GGG CAT TTT TAC CTT GCC AAG TAC TAT GAC AAA TTG ATG CCC ATG GTC Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met Val 1875 1880 1885	5664
ACA GAC AAC AAA ATG GAA AAG CAA GGT GAT CTC ATC CGG TAT ATA GTT Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val 1890 1895 1900	5712
CTT CAT TTT GGC AGA TCT CTA CAA TAT GGA AAT CAG TTC ATA TAT CAG Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln 1905 1910 1915 1920	5760
TCA ATG CCA CGA ATG TTA ACT CTA TGG CTT GAT TAT GGT ACA AAG GCA Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala 1925 1930 1935	5808
TAT GAA TGG GAA AAA GCT GGC CGC TCC GAT CGT GTA CAA ATG AGG AAT Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg Asn 1940 1945 1950	5856
GAT TTG GGT AAA ATA AAC AAG GTT ATC ACA GAG CAT ACA AAC TAT TTA Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr Leu 1955 1960 1965	5904
GCT CCA TAT CAA TTT TTG ACT GCT TTT TCA CAA TTG ATC TCT CGA ATT Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile 1970 1975 1980	5952
TGT CAT TCT CAC GAT GAA GTT TTT GTT GTG CTT GAT GGA AAT AAT AGC Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn Ser 1985 1990 1995 2000	6000
CAA GTA TTT CTA GCC TAT CCT CAA CAA GCA ATG TGG ATG ATG ACA GCT Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr Ala 2005 2010 2015	6048
GTG TCA AAG TCA TCT TAT CCC ATG CGT GTG AAC AGA TGC AAG GAA ATC Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu Ile 2020 2025 2030	6096
CTC AAT AAA GCT ATT CAT ATG AAA AAA TCC TTA GAG AAG TTT GTT GGA Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val Gly 2035 2040 2045	6144
GAT GCA ACT CGC CTA ACA GAT AAG CTT CTA GAA TTG TGC AAT AAA CCG Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro 2050 2055 2060	6192
GTT GAT GGA AGT AGT TCC ACA TTA AGC ATG AGC ACT CAT TTT AAA ATG Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys Met 2065 2070 2075 2080	6240
CTT AAA AAG CTG GTA GAA GAA GCA ACA TTT AGT GAA ATC CTC ATT CCT Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile Pro 2085 2090 2095	6288
CTA CAA TCA GTC ATG ATA CCT ACA CTT CCA TCA ATT CTG GGT ACC CAT Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr His 2100 2105 2110	6336
GCT AAC CAT GCT AGC CAT GAA CCA TTT CCT GGA CAT TGG GCC TAT ATT Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr Ile 2115 2120 2125	6384

-87-

GCA GGG TTT GAT GAT ATG GTG GAA ATT CTT GCT TCT CTT CAG AAA CCA Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys Pro 2130 2135 2140	6432
AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC TAC ATC ATG ATG Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met 2145 2150 2155 2160	6480
TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA CTA ATG GAA TTC Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe 2165 2170 2175	6528
AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG TCT CGT AGA Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg 2180 2185 2190	6576
AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT GAA Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu 2195 2200 2205	6624
TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG AGA CCT ATT Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile 2210 2215 2220	6672
CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAA Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2225 2230 2235 2240	6720
CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2245 2250 2255	6768
AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 2265 2270	6816
GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 2280 2285	6864
AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 2295 2300	6912
ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2305 2310 2315 2320	6960
TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2325 2330 2335	7008
GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 2345 2350	7056
AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg 2355 2360 2365	7104
AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT CAG CGA GAG CCT Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro 2370 2375 2380	7152
TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT CTT GTG GAA TGG Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp 2385 2390 2395 2400	7200

-88-

AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG AAT GAA ACT GGA Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly 2405 2410 2415	7248
GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT GAC ATT GAG CAG Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln 2420 2425 2430	7296
CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG ACA GGA CTG CCG Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro 2435 2440 2445	7344
TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA GAA GCT ACT GAT Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr Asp 2450 2455 2460	7392
GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT CCA TAT ATG Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met 2465 2470 2475	7437
TGA	7440

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2479 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser 1 5 10 15
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu 20 25 30
Met Ser Met Gln Lys Leu Glu Phe Ile Glu Val Thr Leu Leu Thr Val 35 40 45
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu 50 55 60
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile 65 70 75 80
Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly 85 90 95
Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 100 105 110
Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu 115 120 125
Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu 130 135 140
Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys 145 150 155 160
Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp 165 170 175

-89-

Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe
 180 185 190
 Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val
 195 200 205
 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp
 210 215 220
 Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser
 225 230 235 240
 Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu
 245 250 255
 Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350
 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Thr Val Leu Lys Ser Cys Arg
 370 375 380
 Arg Leu Leu Glu Ser Val Gln Lys Arg Thr Gly Gly Asn Ile Asp Lys
 385 390 395 400
 Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser Ser
 405 410 415
 Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu Pro
 420 425 430
 Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr Phe
 435 440 445
 Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr Ser
 450 455 460
 Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro Arg
 465 470 475 480
 Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu Gln
 485 490 495
 Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe Ile
 500 505 510
 Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu Ile
 515 520 525

-90-

Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala Ser
 530 535 540
 Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu Thr
 545 550 555 560
 Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu Phe
 565 570 575
 Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser Gln
 580 585 590
 Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys Lys
 595 600 605
 Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His Leu
 610 615 620
 Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala Val
 625 630 635 640
 Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val Arg
 645 650 655
 Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp Ser
 660 665 670
 Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu Ala
 675 680 685
 Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr Leu
 690 695 700
 Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu Val
 705 710 715 720
 Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser Ala
 725 730 735
 Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala Ala
 740 745 750
 Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro Ile
 755 760 765
 Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala Leu
 770 775 780
 Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val Ala
 785 790 795 800
 His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn Val
 805 810 815
 Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln Val
 820 825 830
 Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala Leu
 835 840 845
 Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile Leu
 850 855 860
 Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys Ser
 865 870 875 880

Lys	Asp	Glu	Leu	Glu	Arg	Ala	Leu	His	Tyr	Leu	Lys	Asn	Glu	Thr	Glu
				885					890						895
Ile	Glu	Leu	Gly	Ser	Leu	Leu	Arg	Gln	Asp	Phe	Gln	Gly	Leu	His	Asn
			900					905					910		
Glu	Leu	Leu	Leu	Arg	Ile	Gly	Glu	His	Tyr	Gln	Gln	Val	Phe	Asn	Gly
		915					920					925			
Leu	Ser	Ile	Leu	Ala	Ser	Phe	Ala	Ser	Ser	Asp	Asp	Pro	Tyr	Gln	Gly
	930					935					940				
Pro	Arg	Asp	Ile	Ile	Ser	Pro	Glu	Leu	Met	Ala	Asp	Tyr	Leu	Gln	Pro
945					950					955					960
Lys	Leu	Leu	Gly	Ile	Leu	Ala	Phe	Phe	Asn	Met	Gln	Leu	Leu	Ser	Ser
				965					970					975	
Ser	Val	Gly	Ile	Glu	Asp	Lys	Lys	Met	Ala	Leu	Asn	Ser	Leu	Met	Ser
			980					985					990		
Leu	Met	Lys	Leu	Met	Gly	Pro	Lys	His	Val	Ser	Ser	Val	Arg	Val	Lys
		995					1000					1005			
Met	Met	Thr	Thr	Leu	Arg	Thr	Gly	Leu	Arg	Phe	Lys	Asp	Asp	Phe	Pro
	1010					1015					1020				
Glu	Leu	Cys	Cys	Arg	Ala	Trp	Asp	Cys	Phe	Val	Arg	Cys	Leu	Asp	His
1025					1030					1035					1040
Ala	Cys	Leu	Gly	Ser	Leu	Leu	Ser	His	Val	Ile	Val	Ala	Leu	Leu	Pro
				1045					1050					1055	
Leu	Ile	His	Ile	Gln	Pro	Lys	Glu	Thr	Ala	Ala	Ile	Phe	His	Tyr	Leu
			1060					1065					1070		
Ile	Ile	Glu	Asn	Arg	Asp	Ala	Val	Gln	Asp	Phe	Leu	His	Glu	Ile	Tyr
		1075					1080					1085			
Phe	Leu	Pro	Asp	His	Pro	Glu	Leu	Lys	Lys	Ile	Lys	Ala	Val	Leu	Gln
	1090					1095					1100				
Glu	Tyr	Arg	Lys	Glu	Thr	Ser	Glu	Ser	Thr	Asp	Leu	Gln	Thr	Thr	Leu
1105					1110					1115					1120
Gln	Leu	Ser	Met	Lys	Ala	Ile	Gln	His	Glu	Asn	Val	Asp	Val	Arg	Ile
				1125					1130					1135	
His	Ala	Leu	Thr	Ser	Leu	Lys	Glu	Thr	Leu	Tyr	Lys	Asn	Gln	Glu	Lys
			1140					1145					1150		
Leu	Ile	Lys	Tyr	Ala	Thr	Asp	Ser	Glu	Thr	Val	Glu	Pro	Ile	Ile	Ser
	1155					1160						1165			
Gln	Leu	Val	Thr	Val	Leu	Leu	Lys	Gly	Cys	Gln	Asp	Ala	Asn	Ser	Gln
	1170					1175					1180				
Ala	Arg	Leu	Leu	Cys	Gly	Glu	Cys	Leu	Gly	Glu	Leu	Gly	Ala	Ile	Asp
1185					1190					1195					1200
Pro	Gly	Arg	Leu	Asp	Phe	Ser	Thr	Thr	Glu	Thr	Gln	Gly	Lys	Asp	Phe
				1205					1210			</			

-92-

Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala
1235 1240 1245

Pro Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp
1250 1255 1260

Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg Arg
1265 1270 1275 1280

Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr Arg
1285 1290 1295

Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys Pro
1300 1305 1310

Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser
1315 1320 1325

Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser Lys
1330 1335 1340

Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val Thr
1345 1350 1355 1360

Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys Asn
1365 1370 1375

Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu Lys
1380 1385 1390

His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu
1395 1400 1405

Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Thr
1410 1415 1420

Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro
1425 1430 1435 1440

His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr Val
1445 1450 1455

Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile Pro
1460 1465 1470

Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg
1475 1480 1485

Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile
1490 1495 1500

Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His Glu
1505 1510 1515 1520

Pro Asp Gly Val Ser Gly Val Ser Ala Ile Arg Lys Ala Glu Pro Ser
1525 1530 1535

Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg Asp
1540 1545 1550

Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile
1555 1560 1565

Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln Leu
1570 1575 1580

-93-

Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser Glu
 1585 1590 1595 1600
 Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu
 1605 1610 1615
 Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser
 1620 1625 1630
 Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys
 1635 1640 1645
 Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu
 1650 1655 1660
 Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln
 1665 1670 1675 1680
 Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu Glu
 1685 1690 1695
 His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln
 1700 1705 1710
 Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser
 1715 1720 1725
 Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser
 1730 1735 1740
 Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu
 1745 1750 1755 1760
 Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr
 1765 1770 1775
 Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val
 1780 1785 1790
 Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu
 1795 1800 1805
 Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu Thr
 1810 1815 1820
 Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu
 1825 1830 1835 1840
 Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile
 1845 1850 1855
 Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu Asp
 1860 1865 1870
 Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met Val
 1875 1880 1885
 Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val
 1890 1895 1900
 Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln
 1905 1910 1915 1920
 Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala
 1925 1930 1935

-94-

Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg Asn
 1940 1945 1950
 Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr Leu
 1955 1960 1965
 Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile
 1970 1975 1980
 Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn Ser
 1985 1990 1995 2000
 Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr Ala
 2005 2010 2015
 Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu Ile
 2020 2025 2030
 Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val Gly
 2035 2040 2045
 Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro
 2050 2055 2060
 Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys Met
 2065 2070 2075 2080
 Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile Pro
 2085 2090 2095
 Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr His
 2100 2105 2110
 Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr Ile
 2115 2120 2125
 Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys Pro
 2130 2135 2140
 Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met
 2145 2150 2155 2160
 Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe
 2165 2170 2175
 Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg
 2180 2185 2190
 Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu
 2195 2200 2205
 Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile
 2210 2215 2220
 Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu
 2225 2230 2235 2240
 Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu
 2245 2250 2255
 Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His
 2260 2265 2270
 Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser
 2275 2280 2285

-95-

Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr
 2290 2295 2300
 Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser
 2305 2310 2315 2320
 Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys
 2325 2330 2335
 Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His
 2340 2345 2350
 Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg
 2355 2360 2365
 Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro
 2370 2375 2380
 Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp
 2385 2390 2395 2400
 Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly
 2405 2410 2415
 Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln
 2420 2425 2430
 Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro
 2435 2440 2445
 Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr Asp
 2450 2455 2460
 Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2465 2470 2475

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vii) IMMEDIATE SOURCE:
- (B) CLONE: Primer oDH26

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGGTTTCTGA GAACATTCCC TGA

23

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- (vii) IMMEDIATE SOURCE:
- (B) CLONE: FLAG tag

-96-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Asp Tyr Lys Asp Asp Asp Asp Lys
 1 5

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

(B) CLONE: Primer 279-3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGGATGATGA CAGCTGTGTC

20

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

(B) CLONE: Primer 279-6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGTAGTCGCT GCTCAATGTC

20

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7624 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 333..7562

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTTGTGAAGA GAATGTTTTA CACTCTTGTT AGTGAAGTTT ATTCTTTAAA AGTCAATCGT

60

CAAGGATTTA GCAAATGAAT TAGCACTTCG GATATACTTG TTTATTTAAT ATCTTTTTTG

120

-97-

TTTATTTCAA AGAATTCAGT AATTGGATCA TAACGAGACT TCTGCGGATT GCAGCAACTC	180
CCTCCTGTCA TTTGTTACAC AAGAAAATCT GTGAAGTCAT CTGTTTCATTA TTATTTCTTT	240
TTAAAAGCAA GAGTCCTGCT ATTTTGGGG TACTCACAAA AGAATTATTA CAACTTTTTG	300
AAGACTTGGT TTACCTCCAT AGAAGAAATG TG ATG GGT CAT GCT GTG GAA TGG	353
Met Gly His Ala Val Glu Trp	
1 5	
CCA GTG GTC ATG AGC CGA TTT TTA AGT CAA TTA GAT GAA CAC ATG GGA	401
Pro Val Val Met Ser Arg Phe Leu Ser Gln Leu Asp Glu His Met Gly	
10 15 20	
TAT TTA CAA TCA GCT CCT TTG CAG TTG ATG AGT ATG CAA AAT TTA GAA	449
Tyr Leu Gln Ser Ala Pro Leu Gln Leu Met Ser Met Gln Asn Leu Glu	
25 30 35	
TTT ATT GAA GTC ACT TTA TTA ATG GTT CTT ACT CGT ATT ATT GCA ATT	497
Phe Ile Glu Val Thr Leu Leu Met Val Leu Thr Arg Ile Ile Ala Ile	
40 45 50 55	
GTG TTT TTT AGA AGG CAA GAA CTC TTA CTT TGG CAG ATA GGT TGT GTT	545
Val Phe Phe Arg Arg Gln Glu Leu Leu Leu Trp Gln Ile Gly Cys Val	
60 65 70	
CTG CTA GAG TAT GGT AGT CCA AAA ATT AAA TCC CTA GCA ATT AGC TTT	593
Leu Leu Glu Tyr Gly Ser Pro Lys Ile Lys Ser Leu Ala Ile Ser Phe	
75 80 85	
TTA ACA GAA CTT TTT CAG CTT GGA GGA CTA CCA GCA CAA CCA GCT AGC	641
Leu Thr Glu Leu Phe Gln Leu Gly Gly Leu Pro Ala Gln Pro Ala Ser	
90 95 100	
ACT TTT TTC AGC TCA TTT TTG GAA TTA TTA AAA CAC CTT GTA GAA ATG	689
Thr Phe Phe Ser Ser Phe Leu Glu Leu Leu Lys His Leu Val Glu Met	
105 110 115	
GAT ACT GAC CAA TTG AAA CTC TAT GAA GAG CCA TTA TCA AAG CTG ATA	737
Asp Thr Asp Gln Leu Lys Leu Tyr Glu Glu Pro Leu Ser Lys Leu Ile	
120 125 130 135	
AAG ACA CTA TTT CCC TTT GAA GCA GAA GCT TAT AGA AAT ATT GAA CCT	785
Lys Thr Leu Phe Pro Phe Glu Ala Glu Ala Tyr Arg Asn Ile Glu Pro	
140 145 150	
GTC TAT TTA AAT ATG CTG CTG GAA AAA CTC TGT GTC ATG TTT GAA GAC	833
Val Tyr Leu Asn Met Leu Leu Glu Lys Leu Cys Val Met Phe Glu Asp	
155 160 165	
GGT GTG CTC ATG CGG CTT AAG TCT GAT TTG CTA AAA GCA GCT TTG TGC	881
Gly Val Leu Met Arg Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys	
170 175 180	
CAT TTA CTG CAG TAT TTC CTT AAA TTT GTG CCA GCT GGG TAT GAA TCT	929
His Leu Leu Gln Tyr Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser	
185 190 195	
GCT TTA CAA GTC AGG AAG GTC TAT GTG AGA AAT ATT TGT AAA GCT CTT	977
Ala Leu Gln Val Arg Lys Val Tyr Val Arg Asn Ile Cys Lys Ala Leu	
200 205 210 215	
TTG GAT GTG CTT GGA ATT GAG GTA GAT GCA GAG TAC TTG TTG GGC CCA	1025
Leu Asp Val Leu Gly Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro	
220 225 230	

-98-

CTT TAT GCA GCT TTG AAA ATG GAA AGT ATG GAA ATC ATT GAG GAG ATT Leu Tyr Ala Ala Leu Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile 235 240 245	1073
CAA TGC CAA ACT CAA CAG GAA AAC CTC AGC AGT AAT AGT GAT GGA ATA Gln Cys Gln Thr Gln Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile 250 255 260	1121
TCA CCC AAA AGG CGT CGT CTC AGC TCG TCT CTA AAC CCT TCT AAA AGA Ser Pro Lys Arg Arg Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg 265 270 275	1169
GCA CCA AAA CAG ACT GAG GAA ATT AAA CAT GTG GAC ATG AAC CAA AAG Ala Pro Lys Gln Thr Glu Glu Ile Lys His Val Asp Met Asn Gln Lys 280 285 290 295	1217
AGC ATA TTA TGG AGT GCA CTG AAA CAG AAA GCT GAA TCC CTT CAG ATT Ser Ile Leu Trp Ser Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile 300 305 310	1265
TCC CTT GAA TAC AGT GGC CTA AAG AAT CCT GTT ATT GAG ATG TTA GAA Ser Leu Glu Tyr Ser Gly Leu Lys Asn Pro Val Ile Glu Met Leu Glu 315 320 325	1313
GGA ATT GCT GTT GTC TTA CAA CTG ACT GCT CTG TGT ACT GTT CAT TGT Gly Ile Ala Val Val Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys 330 335 340	1361
TCT CAT CAA AAC ATG AAC TGC CGT ACT TTC AAG GAC TGT CAA CAT AAA Ser His Gln Asn Met Asn Cys Arg Thr Phe Lys Asp Cys Gln His Lys 345 350 355	1409
TCC AAG AAG AAA CCT TCT GTA GTG ATA ACT TGG ATG TCA TTG GAT TTT Ser Lys Lys Lys Pro Ser Val Val Ile Thr Trp Met Ser Leu Asp Phe 360 365 370 375	1457
TAC ACA AAA GTG CTT AAG AGC TGT AGA AGT TTG TTA GAA TCT GTT CAG Tyr Thr Lys Val Leu Lys Ser Cys Arg Ser Leu Leu Glu Ser Val Gln 380 385 390	1505
AAA CTG GAC CTG GAG GCA ACC ATT GAT AAG GTG GTG AAA ATT TAT GAT Lys Leu Asp Leu Glu Ala Thr Ile Asp Lys Val Val Lys Ile Tyr Asp 395 400 405	1553
GCT TTG ATT TAT ATG CAA GTA AAC AGT TCA TTT GAA GAT CAT ATC CTG Ala Leu Ile Tyr Met Gln Val Asn Ser Ser Phe Glu Asp His Ile Leu 410 415 420	1601
GAA GAT TTA TGT GGA ATG CTC TCA CTT CCA TGG ATT TAT TCC CAT TCT Glu Asp Leu Cys Gly Met Leu Ser Leu Pro Trp Ile Tyr Ser His Ser 425 430 435	1649
GAT GAT GGC TGT TTA AAG TTG ACC ACA TTT GCC GCT AAT CTT CTA ACA Asp Asp Gly Cys Leu Lys Leu Thr Thr Phe Ala Ala Asn Leu Leu Thr 440 445 450 455	1697
TTA AGC TGT AGG ATT TCA GAT AGC TAT TCA CCA CAG GCA CAA TCA CGA Leu Ser Cys Arg Ile Ser Asp Ser Tyr Ser Pro Gln Ala Gln Ser Arg 460 465 470	1745
TGT GTG TTT CTT CTG ACT CTG TTT CCA AGA AGA ATA TTC CTT GAG TGG Cys Val Phe Leu Thr Leu Phe Pro Arg Arg Ile Phe Leu Glu Trp 475 480 485	1793
AGA ACA GCA GTT TAC AAC TGG GCC CTG CAG AGC TCC CAT GAA GTA ATC Arg Thr Ala Val Tyr Asn Trp Ala Leu Gln Ser Ser His Glu Val Ile 490 495 500	1841

-99-

CGG Arg 505	GCT Ala	AGT Ser	TGT Cys	GTT Val	AGT Ser	GGA Gly 510	TTT Phe	TTT Phe	ATC Ile	TTA Leu	TTG Leu 515	CAG Gln	CAG Gln	CAG Gln	AAT Asn	1889
TCT Ser 520	TGT Cys	AAC Asn	AGA Arg	GTT Val	CCC Pro 525	AAG Lys	ATT Ile	CTT Leu	ATA Ile	GAT Asp 530	AAA Lys	GTC Val	AAA Lys	GAT Asp	GAT Asp 535	1937
TCT Ser	GAC Asp	ATT Ile	GTC Val	AAG Lys 540	AAA Lys	GAA Glu	TTT Phe	GCT Ala	TCT Ser 545	ATA Ile	CTT Leu	GGT Gly	CAA Gln	CTT Leu 550	GTC Val	1985
TGT Cys	ACT Thr	CTT Leu	CAC His 555	GGC Gly	ATG Met	TTT Phe	TAT Tyr	CTG Leu 560	ACA Thr	AGT Ser	TCT Ser	TTA Leu	ACA Thr 565	GAA Glu	CCT Pro	2033
TTC Phe	TCT Ser	GAA Glu 570	CAC His	GGA Gly	CAT His	GTG Val	GAC Asp 575	CTC Leu	TTC Phe	TGT Cys	AGG Arg	AAC Asn 580	TTG Leu	AAA Lys	GCC Ala	2081
ACT Thr 585	TCT Ser	CAA Gln	CAT His	GAA Glu	TGT Cys	TCA Ser 590	TCT Ser	TCT Ser	CAA Gln	CTA Leu	AAA Lys 595	GCT Ala	TCT Ser	GTC Val	TGC Cys	2129
AAG Lys 600	CCA Pro	TTC Phe	CTT Leu	TTC Phe	CTA Leu 605	CTG Leu	AAA Lys	AAA Lys	AAA Lys	ATA Ile 610	CCT Pro	AGT Ser	CCA Pro	GTA Val	AAA Lys 615	2177
CTT Leu	GCT Ala	TTC Phe	ATA Ile	GAT Asp 620	AAT Asn	CTA Leu	CAT His	CAT His	CTT Leu 625	TGT Cys	AAG Lys	CAT His	CTT Leu	GAT Asp 630	TTT Phe	2225
AGA Arg	GAA Glu	GAT Asp 635	GAA Glu	ACA Thr	GAT Asp	GTA Val	AAA Lys	GCA Ala 640	GTT Val	CTT Leu	GGA Gly	ACT Thr	TTA Leu 645	TTA Leu	AAT Asn	2273
TTA Leu	ATG Met	GAA Glu 650	GAT Asp	CCA Pro	GAC Asp	AAA Lys	GAT Asp 655	GTT Val	AGA Arg	GTG Val	GCT Ala	TTT Phe 660	AGT Ser	GGA Gly	AAT Asn	2321
ATC Ile 665	AAG Lys	CAC His	ATA Ile	TTG Leu	GAA Glu	TCC Ser 670	TTG Leu	GAC Asp	TCT Ser	GAA Glu	GAT Asp 675	GGA Gly	TTT Phe	ATA Ile	AAG Lys	2369
GAG Glu 680	CTT Leu	TTT Phe	GTC Val	TTA Leu	AGA Arg 685	ATG Met	AAG Lys	GAA Glu	GCA Ala	TAT Tyr 690	ACA Thr	CAT His	GCC Ala	CAA Gln	ATA Ile 695	2417
TCA Ser	AGA Arg	AAT Asn	AAT Asn	GAG Glu 700	CTG Leu	AAG Lys	GAT Asp	ACC Thr	TTG Leu 705	ATT Ile	CTT Leu	ACA Thr	ACA Thr	GGG Gly 710	GAT Asp	2465
ATT Ile	GGA Gly	AGG Arg	GCC Ala 715	GCA Ala	AAA Lys	GGA Gly	GAT Asp	TTG Leu 720	GTA Val	CCA Pro	TTT Phe	GCA Ala	CTC Leu 725	TTA Leu	CAC His	2513
TTA Leu	TTG Leu	CAT His 730	TGT Cys	TTG Leu	TTA Leu	TCC Ser	AAG Lys 735	TCA Ser	GCA Ala	TCT Ser	GTC Val	TCT Ser	GGA Gly 740	GCA Ala	GCA Ala	2561
TAC Tyr 745	ACA Thr	GAA Glu	ATT Ile	AGA Arg	GCT Ala 750	CTG Leu	GTT Val	GCA Ala	GCT Ala	AAA Lys 755	AGT Ser	GTT Val	AAA Lys	CTG Leu	CAA Gln	2609
AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	ATC Ile	TGT Cys 770	CAG Gln	TTT Phe	TTG Leu	GTA Val	GAA Glu 775	2657

-100-

TCC CTT CAC TCT AGT CAG ATG ACA GCA CTT CCG AAT ACT CCA TGC CAG Ser Leu His Ser Ser Gln Met Thr Ala Leu Pro Asn Thr Pro Cys Gln 780 785 790	2705
AAT GCT GAC GTG CGA AAA CAA GAT GTG GCT CAC CAG AGA GAA ATG GCT Asn Ala Asp Val Arg Lys Gln Asp Val Ala His Gln Arg Glu Met Ala 795 800 805	2753
TTA AAT ACG TTG TCT GAA ATT GCC AAC GTT TTC GAC TTT CCT GAT CTT Leu Asn Thr Leu Ser Glu Ile Ala Asn Val Phe Asp Phe Pro Asp Leu 810 815 820	2801
AAT CGT TTT CTT ACT AGG ACA TTA CAA GTT CTA CTA CCT GAT CTT GCT Asn Arg Phe Leu Thr Arg Thr Leu Gln Val Leu Leu Pro Asp Leu Ala 825 830 835	2849
GCC AAA GCA AGC CCT GCA GCT TCT GCT CTC ATT CGA ACT TTA GGA AAA Ala Lys Ala Ser Pro Ala Ala Ser Ala Leu Ile Arg Thr Leu Gly Lys 840 845 850 855	2897
CAA TTA AAT GTC AAT CGT AGA GAG ATT TTA ATA AAC AAC TTC AAA TAT Gln Leu Asn Val Asn Arg Arg Glu Ile Leu Ile Asn Asn Phe Lys Tyr 860 865 870	2945
ATT TTT TCT CAT TTG GTC TGT TCT TGT TCC AAA GAT GAA TTA GAA CGT Ile Phe Ser His Leu Val Cys Ser Cys Ser Lys Asp Glu Leu Glu Arg 875 880 885	2993
GCC CTT CAT TAT CTG AAG AAT GAA ACA GAA ATT GAA CTG GGG AGC CTG Ala Leu His Tyr Leu Lys Asn Glu Thr Glu Ile Glu Leu Gly Ser Leu 890 895 900	3041
TTG AGA CAA GAT TTC CAA GGA TTG CAT AAT GAA TTA TTG CTG CGT ATT Leu Arg Gln Asp Phe Gln Gly Leu His Asn Glu Leu Leu Leu Arg Ile 905 910 915	3089
GGA GAA CAC TAT CAA CAG GTT TTT AAT GGT TTG TCA ATA CTT GCC TCA Gly Glu His Tyr Gln Gln Val Phe Asn Gly Leu Ser Ile Leu Ala Ser 920 925 930 935	3137
TTT GCA TCC AGT GAT GAT CCA TAT CAG GGC CCG AGA GAT ATC ATA TCA Phe Ala Ser Ser Asp Asp Pro Tyr Gln Gly Pro Arg Asp Ile Ile Ser 940 945 950	3185
CCT GAA CTG ATG GCT GAT TAT TTA CAA CCC AAA TTG TTG GGC ATT TTG Pro Glu Leu Met Ala Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu 955 960 965	3233
GCT TTT TTT AAC ATG CAG TTA CTG AGC TCT AGT GTT GGC ATT GAA GAT Ala Phe Phe Asn Met Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp 970 975 980	3281
AAG AAA ATG GCC TTG AAC AGT TTG ATG TCT TTG ATG AAG TTA ATG GGA Lys Lys Met Ala Leu Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly 985 990 995	3329
CCC AAA CAT GTC AGT TCT GTG AGG GTG AAG ATG ATG ACC ACA CTG AGA Pro Lys His Val Ser Ser Val Arg Val Lys Met Met Thr Thr Leu Arg 1000 1005 1010 1015	3377
ACT GGC CTT CGA TTC AAG GAT GAT TTT CCT GAA TTG TGT TGC AGA GCT Thr Gly Leu Arg Phe Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala 1020 1025 1030	3425
TGG GAC TGC TTT GTT CGC TGC CTG GAT CAT GCT TGT CTG GGC TCC CTT Trp Asp Cys Phe Val Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu 1035 1040 1045	3473

-101-

CTC AGT CAT GTA ATA GTA GCT TTG TTA CCT CTT ATA CAC ATC CAG CCT Leu Ser His Val Ile Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro 1050 1055 1060	3521
AAA GAA ACT GCA GCT ATC TTC CAC TAC CTC ATA ATT GAA AAC AGG GAT Lys Glu Thr Ala Ala Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp 1065 1070 1075	3569
GCT GTG CAA GAT TTT CTT CAT GAA ATA TAT TTT TTA CCT GAT CAT CCA Ala Val Gln Asp Phe Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro 1080 1085 1090 1095	3617
GAA TTA AAA AAG ATA AAA GCC GTT CTC CAG GAA TAC AGA AAG GAG ACC Glu Leu Lys Lys Ile Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr 1100 1105 1110	3665
TCT GAG AGC ACT GAT CTT CAG ACA ACT CTT CAG CTC TCT ATG AAG GCC Ser Glu Ser Thr Asp Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala 1115 1120 1125	3713
ATT CAA CAT GAA AAT GTC GAT GTT CGT ATT CAT GCT CTT ACA AGC TTG Ile Gln His Glu Asn Val Asp Val Arg Ile His Ala Leu Thr Ser Leu 1130 1135 1140	3761
AAG GAA ACC TTG TAT AAA AAT CAG GAA AAA CTG ATA AAG TAT GCA ACA Lys Glu Thr Leu Tyr Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr 1145 1150 1155	3809
GAC AGT GAA ACA GTA GAA CCT ATT ATC TCA CAG TTG GTG ACA GTG CTT Asp Ser Glu Thr Val Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu 1160 1165 1170 1175	3857
TTG AAA GGT TGC CAA GAT GCA AAC TCT CAA GCT CGG TTG CTC TGT GGG Leu Lys Gly Cys Gln Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly 1180 1185 1190	3905
GAA TGT TTA GGG GAA TTG GGG GCG ATA GAT CCA GGT CGA TTA GAT TTC Glu Cys Leu Gly Glu Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe 1195 1200 1205	3953
TCA ACA ACT GAA ACT CAA GGA AAA GAT TTT ACA TTT GTG ACT GGA GTA Ser Thr Thr Glu Thr Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val 1210 1215 1220	4001
GAA GAT TCA AGC TTT GCC TAT GGA TTA TTG ATG GAG CTA ACA AGA GCT Glu Asp Ser Ser Phe Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala 1225 1230 1235	4049
TAC CTT GCG TAT GCT GAT AAT AGC CGA GCT CAA GAT TCA GCT GCC TAT Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala Gln Asp Ser Ala Ala Tyr 1240 1245 1250 1255	4097
GCC ATT CAG GAG TTG CTT TCT ATT TAT GAC TGT AGA GAG ATG GAG ACC Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr 1260 1265 1270	4145
AAC GGC CCA GGT CAC CAA TTG TGG AGG AGA TTT CCT GAG CAT GTT CGG Asn Gly Pro Gly His Gln Leu Trp Arg Arg Phe Pro Glu His Val Arg 1275 1280 1285	4193
GAA ATA CTA GAA CCT CAT CTA AAT ACC AGA TAC AAG AGT TCT CAG AAG Glu Ile Leu Glu Pro His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys 1290 1295 1300	4241
TCA ACC GAT TGG TCT GGA GTA AAG AAG CCA ATT TAC TTA AGT AAA TTG Ser Thr Asp Trp Ser Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu 1305 1310 1315	4289

-102-

GGT AGT AAC TTT GCA GAA TGG TCA GCA TCT TGG GCA GGT TAT CTT ATT Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile 1320 1325 1330 1335	4337
ACA AAG GTT CGA CAT GAT CTT GCC AGT AAA ATT TTC ACC TGC TGT AGC Thr Lys Val Arg His Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser 1340 1345 1350	4385
ATT ATG ATG AAG CAT GAT TTC AAA GTG ACC ATC TAT CTT CTT CCA CAT Ile Met Met Lys His Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His 1355 1360 1365	4433
ATT CTG GTG TAT GTC TTA CTG GGT TGT AAT CAA GAA GAT CAG CAG GAG Ile Leu Val Tyr Val Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu 1370 1375 1380	4481
GTT TAT GCA GAA ATT ATG GCA GTT CTA AAG CAT GAC GAT CAG CAT ACC Val Tyr Ala Glu Ile Met Ala Val Leu Lys His Asp Asp Gln His Thr 1385 1390 1395	4529
ATA AAT ACC CAA GAC ATT GCA TCT GAT CTG TGT CAA CTC AGT ACA CAG Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln 1400 1405 1410 1415	4577
ACT GTG TTC TCC ATG CTT GAC CAT CTC ACA CAG TGG GCA AGG CAC AAA Thr Val Phe Ser Met Leu Asp His Leu Thr Gln Trp Ala Arg His Lys 1420 1425 1430	4625
TTT CAG GCA CTG AAA GCT GAG AAA TGT CCA CAC AGC AAA TCA AAC AGA Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg 1435 1440 1445	4673
AAT AAG GTA GAC TCA ATG GTA TCT ACT GTG GAT TAT GAA GAC TAT CAG Asn Lys Val Asp Ser Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln 1450 1455 1460	4721
AGT GTA ACC CGT TTT CTA GAC CTC ATA CCC CAG GAT ACT CTG GCA GTA Ser Val Thr Arg Phe Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val 1465 1470 1475	4769
GCT TCC TTT CGC TCC AAA GCA TAC ACA CGA GCT GTA ATG CAC TTT GAA Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu 1480 1485 1490 1495	4817
TCA TTT ATT ACA GAA AAG AAG CAA AAT ATT CAG GAA CAT CTT GGA TTT Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile Gln Glu His Leu Gly Phe 1500 1505 1510	4865
TTA CAG AAA TTG TAT GCT GCT ATG CAT GAA CCT GAT GGA GTG GCC GGA Leu Gln Lys Leu Tyr Ala Ala Met His Glu Pro Asp Gly Val Ala Gly 1515 1520 1525	4913
GTC AGT GCA ATT AGA AAG GCA GAA CCA TCT CTA AAA GAA CAG ATC CTT Val Ser Ala Ile Arg Lys Ala Glu Pro Ser Leu Lys Glu Gln Ile Leu 1530 1535 1540	4961
GAA CAT GAA AGC CTT GGC TTG CTG AGG GAT GCC ACT GCT TGT TAT GAC Glu His Glu Ser Leu Gly Leu Leu Arg Asp Ala Thr Ala Cys Tyr Asp 1545 1550 1555	5009
AGG GCT ATT CAG CTA GAA CCA GAC CAG ATC ATT CAT TAC CAT GGT GTA Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile Ile His Tyr His Gly Val 1560 1565 1570 1575	5057
GTA AAG TCC ATG TTA GGT CTT GGT CAG CTG TCT ACT GTT ATC ACT CAG Val Lys Ser Met Leu Gly Leu Gly Gln Leu Ser Thr Val Ile Thr Gln 1580 1585 1590	5105

-103-

GTG AAT GGA GTG CAT GCT AAC AGG TCC GAG TGG ACA GAT GAA TTA AAC	5153
Val Asn Gly Val His Ala Asn Arg Ser Glu Trp Thr Asp Glu Leu Asn	
1595 1600 1605	
ACG TAC AGA GTG GAA GCA GCT TGG AAA TTG TCA CAG TGG GAT TTG GTG	5201
Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val	
1610 1615 1620	
GAA AAC TAT TTG GCA GCA GAT GGA AAA TCT ACA ACA TGG AGT GTC AGA	5249
Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg	
1625 1630 1635	
CTG GGA CAG CTA TTA TTA TCA GCC AAA AAA AGA GAT ATC ACA GCT TTT	5297
Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe	
1640 1645 1650 1655	
TAT GAC TCA CTG AAA CTA GTG AGA GCA GAA CAA ATT GTA CCT CTT TCA	5345
Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser	
1660 1665 1670	
GCT GCA AGC TTT GAA AGA GGC TCC TAC CAA CGA GGA TAT GAA TAT ATT	5393
Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile	
1675 1680 1685	
GTG AGA TTG CAC ATG TTA TGT GAG TTG GAG CAT AGC ATC AAA CCA CTT	5441
Val Arg Leu His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu	
1690 1695 1700	
TTC CAG CAT TCT CCA GGT GAC AGT TCT CAA GAA GAT TCT CTA AAC TGG	5489
Phe Gln His Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp	
1705 1710 1715	
GTA GCT CGA CTA GAA ATG ACC CAG AAT TCC TAC AGA GCC AAG GAG CCT	5537
Val Ala Arg Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro	
1720 1725 1730 1735	
ATC CTG GCT CTC CGG AGG GCT TTA CTA AGC CTC AAC AAA AGA CCA GAT	5585
Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp	
1740 1745 1750	
TAC AAT GAA ATG GTT GGA GAA TGC TGG CTG CAG AGT GCC AGG GTA GCT	5633
Tyr Asn Glu Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala	
1755 1760 1765	
AGA AAG GCT GGT CAC CAC CAG ACA GCC TAC AAT GCT CTC CTT AAT GCA	5681
Arg Lys Ala Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala	
1770 1775 1780	
GGG GAA TCA CGA CTC GCT GAA CTG TAC GTG GAA AGG GCA AAG TGG CTC	5729
Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu	
1785 1790 1795	
TGG TCC AAG GGT GAT GTT CAC CAG GCA CTA ATT GTT CTT CAA AAA GGT	5777
Trp Ser Lys Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly	
1800 1805 1810 1815	
GTT GAA TTA TGT TTT CCT GAA AAT GAA ACC CCA CCT GAG GGT AAG AAC	5825
Val Glu Leu Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn	
1820 1825 1830	
ATG TTA ATC CAT GGT CGA GCT ATG CTA CTA GTG GGC CGA TTT ATG GAA	5873
Met Leu Ile His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu	
1835 1840 1845	
GAA ACA GCT AAC TTT GAA AGC AAT GCA ATT ATG AAA AAA TAT AAG GAT	5921
Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp	
1850 1855 1860	

-104-

GTG ACC GCG TGC CTG CCA GAA TGG GAG GAT GGG CAT TTT TAC CTT GCC Val Thr Ala Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala 1865 1870 1875	5969
AAG TAC TAT GAC AAA TTG ATG CCC ATG GTC ACA GAC AAC AAA ATG GAA Lys Tyr Tyr Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu 1880 1885 1890 1895	6017
AAG CAA GGT GAT CTC ATC CGG TAT ATA GTT CTT CAT TTT GGC AGA TCT Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser 1900 1905 1910	6065
CTA CAA TAT GGA AAT CAG TTC ATA TAT CAG TCA ATG CCA CGA ATG TTA Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu 1915 1920 1925	6113
ACT CTA TGG CTT GAT TAT GGT ACA AAG GCA TAT GAA TGG GAA AAA GCT Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala 1930 1935 1940	6161
GGC CGC TCC GAT CGT GTA CAA ATG AGG AAT GAT TTG GGT AAA ATA AAC Gly Arg Ser Asp Arg Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn 1945 1950 1955	6209
AAG GTT ATC ACA GAG CAT ACA AAC TAT TTA GCT CCA TAT CAA TTT TTG Lys Val Ile Thr Glu His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu 1960 1965 1970 1975	6257
ACT GCT TTT TCA CAA TTG ATC TCT CGA ATT TGT CAT TCT CAC GAT GAA Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile Cys His Ser His Asp Glu 1980 1985 1990	6305
GTT TTT GTT GTG CTT GAT GGA AAT AAT AGC CAA GTA TTT CTA GCC TAT Val Phe Val Val Leu Asp Gly Asn Asn Ser Gln Val Phe Leu Ala Tyr 1995 2000 2005	6353
CCT CAA CAA GCA ATG TGG ATG ATG ACA GCT GTG TCA AAG TCA TCT TAT Pro Gln Gln Ala Met Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr 2010 2015 2020	6401
CCC ATG CGT GTG AAC AGA TGC AAG GAA ATC CTC AAT AAA GCT ATT CAT Pro Met Arg Val Asn Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His 2025 2030 2035	6449
ATG AAA AAA TCC TTA GAG AAG TTT GTT GGA GAT GCA ACT CGC CTA ACA Met Lys Lys Ser Leu Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr 2040 2045 2050 2055	6497
GAT AAG CTT CTA GAA TTG TGC AAT AAA CCG GTG GAA ATT CTT GCT TCT Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro Val Glu Ile Leu Ala Ser 2060 2065 2070	6545
CTT CAG AAA CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe 2075 2080 2085	6593
TAC ATC ATG ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg 2090 2095 2100	6641
CTA ATG GAA TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala 2105 2110 2115	6689
GAG TCT CGT AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA Glu Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro 2120 2125 2130 2135	6737

-105-

CTA AAT GAT GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly 2140 2145 2150	6785
TTG AGA CCT ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met 2155 2160 2165	6833
ACA GGA AAA GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu 2170 2175 2180	6881
TCT GAA AAA CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro 2185 2190 2195	6929
CCT ATT TTT CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser 2200 2205 2210 2215	6977
TGG TAC AGT AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser 2220 2225 2230	7025
ATG GTT GGT TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile 2235 2240 2245	7073
CTC TTT GAT TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys 2250 2255 2260	7121
CTT TTC AAT AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe 2265 2270 2275	7169
CGC CTG ACT CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu 2280 2285 2290 2295	7217
GGT CTT TTT CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp 2300 2305 2310	7265
CAG CGA GAG CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro 2315 2320 2325	7313
CTT GTG GAA TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu 2330 2335 2340	7361
AAT GAA ACT GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu 2345 2350 2355	7409
GAC ATT GAG CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val 2360 2365 2370 2375	7457
ACA GGA CTG CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln 2380 2385 2390	7505
GAA GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr 2395 2400 2405	7553

-106-

CCA TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA
 Pro Tyr Met
 2410

7602

AGTAAAAAAA AAAAAAAAAA AA

7624

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2410 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met	Gly	His	Ala	Val	Glu	Trp	Pro	Val	Val	Met	Ser	Arg	Phe	Leu	Ser	1	5	10	15
Gln	Leu	Asp	Glu	His	Met	Gly	Tyr	Leu	Gln	Ser	Ala	Pro	Leu	Gln	Leu	20	25	30	
Met	Ser	Met	Gln	Asn	Leu	Glu	Phe	Ile	Glu	Val	Thr	Leu	Leu	Met	Val	35	40	45	
Leu	Thr	Arg	Ile	Ile	Ala	Ile	Val	Phe	Phe	Arg	Arg	Gln	Glu	Leu	Leu	50	55	60	
Leu	Trp	Gln	Ile	Gly	Cys	Val	Leu	Leu	Glu	Tyr	Gly	Ser	Pro	Lys	Ile	65	70	75	80
Lys	Ser	Leu	Ala	Ile	Ser	Phe	Leu	Thr	Glu	Leu	Phe	Gln	Leu	Gly	Gly	85	90	95	
Leu	Pro	Ala	Gln	Pro	Ala	Ser	Thr	Phe	Phe	Ser	Ser	Phe	Leu	Glu	Leu	100	105	110	
Leu	Lys	His	Leu	Val	Glu	Met	Asp	Thr	Asp	Gln	Leu	Lys	Leu	Tyr	Glu	115	120	125	
Glu	Pro	Leu	Ser	Lys	Leu	Ile	Lys	Thr	Leu	Phe	Pro	Phe	Glu	Ala	Glu	130	135	140	
Ala	Tyr	Arg	Asn	Ile	Glu	Pro	Val	Tyr	Leu	Asn	Met	Leu	Leu	Glu	Lys	145	150	155	160
Leu	Cys	Val	Met	Phe	Glu	Asp	Gly	Val	Leu	Met	Arg	Leu	Lys	Ser	Asp	165	170	175	
Leu	Leu	Lys	Ala	Ala	Leu	Cys	His	Leu	Leu	Gln	Tyr	Phe	Leu	Lys	Phe	180	185	190	
Val	Pro	Ala	Gly	Tyr	Glu	Ser	Ala	Leu	Gln	Val	Arg	Lys	Val	Tyr	Val	195	200	205	
Arg	Asn	Ile	Cys	Lys	Ala	Leu	Leu	Asp	Val	Leu	Gly	Ile	Glu	Val	Asp	210	215	220	
Ala	Glu	Tyr	Leu	Leu	Gly	Pro	Leu	Tyr	Ala	Ala	Leu	Lys	Met	Glu	Ser	225	230	235	240
Met	Glu	Ile	Ile	Glu	Glu	Ile	Gln	Cys	Gln	Thr	Gln	Gln	Glu	Asn	Leu	245	250	255	

-107-

Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350
 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg
 370 375 380
 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp
 385 390 395 400
 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser
 405 410 415
 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu
 420 425 430
 Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr
 435 440 445
 Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr
 450 455 460
 Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro
 465 470 475 480
 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu
 485 490 495
 Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe
 500 505 510
 Ile Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu
 515 520 525
 Ile Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala
 530 535 540
 Ser Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu
 545 550 555 560
 Thr Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu
 565 570 575
 Phe Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser
 580 585 590
 Gln Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys
 595 600 605

-108-

Lys Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His
 610 615 620
 Leu Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala
 625 630 635 640
 Val Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val
 645 650 655
 Arg Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp
 660 665 670
 Ser Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu
 675 680 685
 Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr
 690 695 700
 Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu
 705 710 715 720
 Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser
 725 730 735
 Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala
 740 745 750
 Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro
 755 760 765
 Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala
 770 775 780
 Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val
 785 790 795 800
 Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn
 805 810 815
 Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln
 820 825 830
 Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala
 835 840 845
 Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile
 850 855 860
 Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys
 865 870 875 880
 Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr
 885 890 895
 Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His
 900 905 910
 Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn
 915 920 925
 Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln
 930 935 940
 Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln
 945 950 955 960

-109-

Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser
 965 970 975
 Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met
 980 985 990
 Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val
 995 1000 1005
 Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe
 1010 1015 1020
 Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp
 1025 1030 1035 1040
 His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu
 1045 1050 1055
 Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr
 1060 1065 1070
 Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile
 1075 1080 1085
 Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu
 1090 1095 1100
 Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr
 1105 1110 1115 1120
 Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg
 1125 1130 1135
 Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu
 1140 1145 1150
 Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile
 1155 1160 1165
 Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser
 1170 1175 1180
 Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile
 1185 1190 1195 1200
 Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp
 1205 1210 1215
 Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu
 1220 1225 1230
 Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg
 1235 1240 1245
 Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr
 1250 1255 1260
 Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg
 1265 1270 1275 1280
 Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr
 1285 1290 1295
 Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys
 1300 1305 1310

-110-

Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala
 1315 1320 1325
 Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser
 1330 1335 1340
 Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val
 1345 1350 1355 1360
 Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys
 1365 1370 1375
 Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu
 1380 1385 1390
 Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp
 1395 1400 1405
 Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu
 1410 1415 1420
 Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys
 1425 1430 1435 1440
 Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr
 1445 1450 1455
 Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile
 1460 1465 1470
 Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr
 1475 1480 1485
 Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn
 1490 1495 1500
 Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His
 1505 1510 1515 1520
 Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro
 1525 1530 1535
 Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg
 1540 1545 1550
 Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln
 1555 1560 1565
 Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln
 1570 1575 1580
 Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser
 1585 1590 1595 1600
 Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys
 1605 1610 1615
 Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys
 1620 1625 1630
 Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys
 1635 1640 1645
 Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala
 1650 1655 1660

-111-

Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr
 1665 1670 1675 1680
 Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu
 1685 1690 1695
 Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser
 1700 1705 1710
 Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn
 1715 1720 1725
 Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu
 1730 1735 1740
 Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp
 1745 1750 1755 1760
 Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala
 1765 1770 1775
 Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr
 1780 1785 1790
 Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala
 1795 1800 1805
 Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu
 1810 1815 1820
 Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu
 1825 1830 1835 1840
 Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala
 1845 1850 1855
 Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu
 1860 1865 1870
 Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met
 1875 1880 1885
 Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile
 1890 1895 1900
 Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr
 1905 1910 1915 1920
 Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys
 1925 1930 1935
 Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg
 1940 1945 1950
 Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr
 1955 1960 1965
 Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg
 1970 1975 1980
 Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn
 1985 1990 1995 2000
 Ser Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr
 2005 2010 2015

-112-

Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu
 2020 2025 2030
 Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val
 2035 2040 2045
 Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys
 2050 2055 2060
 Pro Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu
 2065 2070 2075 2080
 Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp
 2085 2090 2095
 Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn
 2100 2105 2110
 Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile
 2115 2120 2125
 Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu
 2130 2135 2140
 Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr
 2145 2150 2155 2160
 Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met
 2165 2170 2175
 Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu
 2180 2185 2190
 Phe Leu Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg
 2195 2200 2205
 Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys
 2210 2215 2220
 Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly
 2225 2230 2235 2240
 Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys
 2245 2250 2255
 Val His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu
 2260 2265 2270
 Val Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly
 2275 2280 2285
 Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val
 2290 2295 2300
 Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu
 2305 2310 2315 2320
 Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys
 2325 2330 2335
 Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu
 2340 2345 2350
 Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val
 2355 2360 2365

-113-

Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly
 2370 2375 2380

His Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys
 2385 2390 2395 2400

Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2405 2410

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7502 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..7440

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATG GGT CAT GCT GTG GAA TGG CCA GTG GTC ATG AGC CGA TTT TTA AGT	48
Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser	
1 5 10 15	
CAA TTA GAT GAA CAC ATG GGA TAT TTA CAA TCA GCT CCT TTG CAG TTG	96
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu	
20 25 30	
ATG AGT ATG CAA AAT TTA GAA TTT ATT GAA GTC ACT TTA TTA ATG GTT	144
Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val	
35 40 45	
CTT ACT CGT ATT ATT GCA ATT GTG TTT TTT AGA AGG CAA GAA CTC TTA	192
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu	
50 55 60	
CTT TGG CAG ATA GGT TGT GTT CTG CTA GAG TAT GGT AGT CCA AAA ATT	240
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile	
65 70 75 80	
AAA TCC CTA GCA ATT AGC TTT TTA ACA GAA CTT TTT CAG CTT GGA GGA	288
Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly	
85 90 95	
CTA CCA GCA CAA CCA GCT AGC ACT TTT TTC AGC TCA TTT TTG GAA TTA	336
Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu	
100 105 110	
TTA AAA CAC CTT GTA GAA ATG GAT ACT GAC CAA TTG AAA CTC TAT GAA	384
Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu	
115 120 125	
GAG CCA TTA TCA AAG CTG ATA AAG ACA CTA TTT CCC TTT GAA GCA GAA	432
Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu	
130 135 140	
GCT TAT AGA AAT ATT GAA CCT GTC TAT TTA AAT ATG CTG CTG GAA AAA	480
Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys	
145 150 155 160	

-114-

CTC	TGT	GTC	ATG	TTT	GAA	GAC	GGT	GTG	CTC	ATG	CGG	CTT	AAG	TCT	GAT	528
Leu	Cys	Val	Met	Phe	Glu	Asp	Gly	Val	Leu	Met	Arg	Leu	Lys	Ser	Asp	
			165					170						175		
TTG	CTA	AAA	GCA	GCT	TTG	TGC	CAT	TTA	CTG	CAG	TAT	TTC	CTT	AAA	TTT	576
Leu	Leu	Lys	Ala	Ala	Leu	Cys	His	Leu	Leu	Gln	Tyr	Phe	Leu	Lys	Phe	
			180					185					190			
GTG	CCA	GCT	GGG	TAT	GAA	TCT	GCT	TTA	CAA	GTC	AGG	AAG	GTC	TAT	GTG	624
Val	Pro	Ala	Gly	Tyr	Glu	Ser	Ala	Leu	Gln	Val	Arg	Lys	Val	Tyr	Val	
			195				200					205				
AGA	AAT	ATT	TGT	AAA	GCT	CTT	TTG	GAT	GTG	CTT	GGA	ATT	GAG	GTA	GAT	672
Arg	Asn	Ile	Cys	Lys	Ala	Leu	Leu	Asp	Val	Leu	Gly	Ile	Glu	Val	Asp	
	210					215					220					
GCA	GAG	TAC	TTG	TTG	GGC	CCA	CTT	TAT	GCA	GCT	TTG	AAA	ATG	GAA	AGT	720
Ala	Glu	Tyr	Leu	Leu	Gly	Pro	Leu	Tyr	Ala	Ala	Leu	Lys	Met	Glu	Ser	
225					230				235						240	
ATG	GAA	ATC	ATT	GAG	GAG	ATT	CAA	TGC	CAA	ACT	CAA	CAG	GAA	AAC	CTC	768
Met	Glu	Ile	Ile	Glu	Glu	Ile	Gln	Cys	Gln	Thr	Gln	Gln	Glu	Asn	Leu	
				245				250						255		
AGC	AGT	AAT	AGT	GAT	GGA	ATA	TCA	CCC	AAA	AGG	CGT	CGT	CTC	AGC	TCG	816
Ser	Ser	Asn	Ser	Asp	Gly	Ile	Ser	Pro	Lys	Arg	Arg	Arg	Leu	Ser	Ser	
			260					265					270			
TCT	CTA	AAC	CCT	TCT	AAA	AGA	GCA	CCA	AAA	CAG	ACT	GAG	GAA	ATT	AAA	864
Ser	Leu	Asn	Pro	Ser	Lys	Arg	Ala	Pro	Lys	Gln	Thr	Glu	Glu	Ile	Lys	
		275					280					285				
CAT	GTG	GAC	ATG	AAC	CAA	AAG	AGC	ATA	TTA	TGG	AGT	GCA	CTG	AAA	CAG	912
His	Val	Asp	Met	Asn	Gln	Lys	Ser	Ile	Leu	Trp	Ser	Ala	Leu	Lys	Gln	
	290					295					300					
AAA	GCT	GAA	TCC	CTT	CAG	ATT	TCC	CTT	GAA	TAC	AGT	GGC	CTA	AAG	AAT	960
Lys	Ala	Glu	Ser	Leu	Gln	Ile	Ser	Leu	Glu	Tyr	Ser	Gly	Leu	Lys	Asn	
305					310					315					320	
CCT	GTT	ATT	GAG	ATG	TTA	GAA	GGA	ATT	GCT	GTT	GTC	TTA	CAA	CTG	ACT	1008
Pro	Val	Ile	Glu	Met	Leu	Glu	Gly	Ile	Ala	Val	Val	Leu	Gln	Leu	Thr	
				325					330					335		
GCT	CTG	TGT	ACT	GTT	CAT	TGT	TCT	CAT	CAA	AAC	ATG	AAC	TGC	CGT	ACT	1056
Ala	Leu	Cys	Thr	Val	His	Cys	Ser	His	Gln	Asn	Met	Asn	Cys	Arg	Thr	
			340					345					350			
TTC	AAG	GAC	TGT	CAA	CAT	AAA	TCC	AAG	AAG	AAA	CCT	TCT	GTA	GTG	ATA	1104
Phe	Lys	Asp	Cys	Gln	His	Lys	Ser	Lys	Lys	Lys	Pro	Ser	Val	Val	Ile	
		355					360					365				
ACT	TGG	ATG	TCA	TTG	GAT	TTT	TAC	ACA	AAA	GTG	CTT	AAG	AGC	TGT	AGA	1152
Thr	Trp	Met	Ser	Leu	Asp	Phe	Tyr	Thr	Lys	Val	Leu	Lys	Ser	Cys	Arg	
	370					375					380					
AGT	TTG	TTA	GAA	TCT	GTT	CAG	AAA	CTG	GAC	CTG	GAG	GCA	ACC	ATT	GAT	1200
Ser	Leu	Leu	Glu	Ser	Val	Gln	Lys	Leu	Asp	Leu	Glu	Ala	Thr	Ile	Asp	
385					390					395					400	
AAG	GTG	GTG	AAA	ATT	TAT	GAT	GCT	TTG	ATT	TAT	ATG	CAA	GTA	AAC	AGT	1248
Lys	Val	Val	Lys	Ile	Tyr	Asp	Ala	Leu	Ile	Tyr	Met	Gln	Val	Asn	Ser	
			405					410						415		
TCA	TTT	GAA	GAT	CAT	ATC	CTG	GAA	GAT	TTA	TGT	GGA	ATG	CTC	TCA	CTT	1296
Ser	Phe	Glu	Asp	His	Ile	Leu	Glu	Asp	Leu	Cys	Gly	Met	Leu	Ser	Leu	
			420					425					430			

-115-

CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 440 445	1344
TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 455 460	1392
TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 465 470 475 480	1440
AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu 485 490 495	1488
CAG AGC TCC CAT GAA GTA ATC CGG GCT AGT TGT GTT AGT GGA TTT TTT Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe 500 505 510	1536
ATC TTA TTG CAG CAG CAG AAT TCT TGT AAC AGA GTT CCC AAG ATT CTT Ile Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu 515 520 525	1584
ATA GAT AAA GTC AAA GAT GAT TCT GAC ATT GTC AAG AAA GAA TTT GCT Ile Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala 530 535 540	1632
TCT ATA CTT GGT CAA CTT GTC TGT ACT CTT CAC GGC ATG TTT TAT CTG Ser Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu 545 550 555 560	1680
ACA AGT TCT TTA ACA GAA CCT TTC TCT GAA CAC GGA CAT GTG GAC CTC Thr Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu 565 570 575	1728
TTC TGT AGG AAC TTG AAA GCC ACT TCT CAA CAT GAA TGT TCA TCT TCT Phe Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser 580 585 590	1776
CAA CTA AAA GCT TCT GTC TGC AAG CCA TTC CTT TTC CTA CTG AAA AAA Gln Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys 595 600 605	1824
AAA ATA CCT AGT CCA GTA AAA CTT GCT TTC ATA GAT AAT CTA CAT CAT Lys Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His 610 615 620	1872
CTT TGT AAG CAT CTT GAT TTT AGA GAA GAT GAA ACA GAT GTA AAA GCA Leu Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala 625 630 635 640	1920
GTT CTT GGA ACT TTA TTA AAT TTA ATG GAA GAT CCA GAC AAA GAT GTT Val Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val 645 650 655	1968
AGA GTG GCT TTT AGT GGA AAT ATC AAG CAC ATA TTG GAA TCC TTG GAC Arg Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp 660 665 670	2016
TCT GAA GAT GGA TTT ATA AAG GAG CTT TTT GTC TTA AGA ATG AAG GAA Ser Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu 675 680 685	2064
GCA TAT ACA CAT GCC CAA ATA TCA AGA AAT AAT GAG CTG AAG GAT ACC Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr 690 695 700	2112

-116-

TTG ATT CTT ACA ACA GGG GAT ATT GGA AGG GCC GCA AAA GGA GAT TTG Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu 705 710 715 720	2160
GTA CCA TTT GCA CTC TTA CAC TTA TTG CAT TGT TTG TTA TCC AAG TCA Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser 725 730 735	2208
GCA TCT GTC TCT GGA GCA GCA TAC ACA GAA ATT AGA GCT CTG GTT GCA Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala 740 745 750	2256
GCT AAA AGT GTT AAA CTG CAA AGT TTT TTC AGC CAG TAT AAG AAA CCC Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro 755 760 765	2304
ATC TGT CAG TTT TTG GTA GAA TCC CTT CAC TCT AGT CAG ATG ACA GCA Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala 770 775 780	2352
CTT CCG AAT ACT CCA TGC CAG AAT GCT GAC GTG CGA AAA CAA GAT GTG Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val 785 790 795 800	2400
GCT CAC CAG AGA GAA ATG GCT TTA AAT ACG TTG TCT GAA ATT GCC AAC Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn 805 810 815	2448
GTT TTC GAC TTT CCT GAT CTT AAT CGT TTT CTT ACT AGG ACA TTA CAA Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln 820 825 830	2496
GTT CTA CTA CCT GAT CTT GCT GCC AAA GCA AGC CCT GCA GCT TCT GCT Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala 835 840 845	2544
CTC ATT CGA ACT TTA GGA AAA CAA TTA AAT GTC AAT CGT AGA GAG ATT Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile 850 855 860	2592
TTA ATA AAC AAC TTC AAA TAT ATT TTT TCT CAT TTG GTC TGT TCT TGT Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys 865 870 875 880	2640
TCC AAA GAT GAA TTA GAA CGT GCC CTT CAT TAT CTG AAG AAT GAA ACA Ser Lys Asp Glu Leu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr 885 890 895	2688
GAA ATT GAA CTG GGG AGC CTG TTG AGA CAA GAT TTC CAA GGA TTG CAT Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His 900 905 910	2736
AAT GAA TTA TTG CTG CGT ATT GGA GAA CAC TAT CAA CAG GTT TTT AAT Asn Glu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn 915 920 925	2784
GGT TTG TCA ATA CTT GCC TCA TTT GCA TCC AGT GAT GAT CCA TAT CAG Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln 930 935 940	2832
GGC CCG AGA GAT ATC ATA TCA CCT GAA CTG ATG GCT GAT TAT TTA CAA Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln 945 950 955 960	2880
CCC AAA TTG TTG GGC ATT TTG GCT TTT TTT AAC ATG CAG TTA CTG AGC Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser 965 970 975	2928

-117-

TCT AGT GTT GGC ATT GAA GAT AAG AAA ATG GCC TTG AAC AGT TTG ATG Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met 980 985 990	2976
TCT TTG ATG AAG TTA ATG GGA CCC AAA CAT GTC AGT TCT GTG AGG GTG Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val 995 1000 1005	3024
AAG ATG ATG ACC ACA CTG AGA ACT GGC CTT CGA TTC AAG GAT GAT TTT Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe 1010 1015 1020	3072
CCT GAA TTG TGT TGC AGA GCT TGG GAC TGC TTT GTT CGC TGC CTG GAT Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp 1025 1030 1035 1040	3120
CAT GCT TGT CTG GGC TCC CTT CTC AGT CAT GTA ATA GTA GCT TTG TTA His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu 1045 1050 1055	3168
CCT CTT ATA CAC ATC CAG CCT AAA GAA ACT GCA GCT ATC TTC CAC TAC Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr 1060 1065 1070	3216
CTC ATA ATT GAA AAC AGG GAT GCT GTG CAA GAT TTT CTT CAT GAA ATA Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile 1075 1080 1085	3264
TAT TTT TTA CCT GAT CAT CCA GAA TTA AAA AAG ATA AAA GCC GTT CTC Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu 1090 1095 1100	3312
CAG GAA TAC AGA AAG GAG ACC TCT GAG AGC ACT GAT CTT CAG ACA ACT Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr 1105 1110 1115 1120	3360
CTT CAG CTC TCT ATG AAG GCC ATT CAA CAT GAA AAT GTC GAT GTT CGT Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg 1125 1130 1135	3408
ATT CAT GCT CTT ACA AGC TTG AAG GAA ACC TTG TAT AAA AAT CAG GAA Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu 1140 1145 1150	3456
AAA CTG ATA AAG TAT GCA ACA GAC AGT GAA ACA GTA GAA CCT ATT ATC Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile 1155 1160 1165	3504
TCA CAG TTG GTG ACA GTG CTT TTG AAA GGT TGC CAA GAT GCA AAC TCT Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser 1170 1175 1180	3552
CAA GCT CGG TTG CTC TGT GGG GAA TGT TTA GGG GAA TTG GGG GCG ATA Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile 1185 1190 1195 1200	3600
GAT CCA GGT CGA TTA GAT TTC TCA ACA ACT GAA ACT CAA GGA AAA GAT Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp 1205 1210 1215	3648
TTT ACA TTT GTG ACT GGA GTA GAA GAT TCA AGC TTT GCC TAT GGA TTA Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu 1220 1225 1230	3696
TTG ATG GAG CTA ACA AGA GCT TAC CTT GCG TAT GCT GAT AAT AGC CGA Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg 1235 1240 1245	3744

-118-

GCT CAA GAT TCA GCT GCC TAT GCC ATT CAG GAG TTG CTT TCT ATT TAT Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr 1250 1255 1260	3792
GAC TGT AGA GAG ATG GAG ACC AAC GGC CCA GGT CAC CAA TTG TGG AGG Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg 1265 1270 1275 1280	3840
AGA TTT CCT GAG CAT GTT CGG GAA ATA CTA GAA CCT CAT CTA AAT ACC Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr 1285 1290 1295	3888
AGA TAC AAG AGT TCT CAG AAG TCA ACC GAT TGG TCT GGA GTA AAG AAG Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys 1300 1305 1310	3936
CCA ATT TAC TTA AGT AAA TTG GGT AGT AAC TTT GCA GAA TGG TCA GCA Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325	3984
TCT TGG GCA GGT TAT CTT ATT ACA AAG GTT CGA CAT GAT CTT GCC AGT Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1335 1340	4032
AAA ATT TTC ACC TGC TGT AGC ATT ATG ATG AAG CAT GAT TTC AAA GTG Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360	4080
ACC ATC TAT CTT CTT CCA CAT ATT CTG GTG TAT GTC TTA CTG GGT TGT Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375	4128
AAT CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT ATG GCA GTT CTA Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390	4176
AAG CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC ATT GCA TCT GAT Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 1400 1405	4224
CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420	4272
ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440	4320
CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455	4368
GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470	4416
CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485	4464
CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500	4512
ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515 1520	4560

-119-

GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535	4608
TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550	4656
GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565	4704
ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1575 1580	4752
CTG TCT ACT GTT ATC ACT CAG GTG AAT GGA GTG CAT GCT AAC AGG TCC Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600	4800
GAG TGG ACA GAT GAA TTA AAC ACG TAC AGA GTG GAA GCA GCT TGG AAA Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615	4848
TTG TCA CAG TGG GAT TTG GTG GAA AAC TAT TTG GCA GCA GAT GGA AAA Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630	4896
TCT ACA ACA TGG AGT GTC AGA CTG GGA CAG CTA TTA TTA TCA GCC AAA Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys 1635 1640 1645	4944
AAA AGA GAT ATC ACA GCT TTT TAT GAC TCA CTG AAA CTA GTG AGA GCA Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660	4992
GAA CAA ATT GTA CCT CTT TCA GCT GCA AGC TTT GAA AGA GGC TCC TAC Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1665 1670 1675 1680	5040
CAA CGA GGA TAT GAA TAT ATT GTG AGA TTG CAC ATG TTA TGT GAG TTG Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 1695	5088
GAG CAT AGC ATC AAA CCA CTT TTC CAG CAT TCT CCA GGT GAC AGT TCT Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser 1700 1705 1710	5136
CAA GAA GAT TCT CTA AAC TGG GTA GCT CGA CTA GAA ATG ACC CAG AAT Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn 1715 1720 1725	5184
TCC TAC AGA GCC AAG GAG CCT ATC CTG GCT CTC CGG AGG GCT TTA CTA Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1735 1740	5232
AGC CTC AAC AAA AGA CCA GAT TAC AAT GAA ATG GTT GGA GAA TGC TGG Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp 1745 1750 1755 1760	5280
CTG CAG AGT GCC AGG GTA GCT AGA AAG GCT GGT CAC CAC CAG ACA GCC Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala 1765 1770 1775	5328
TAC AAT GCT CTC CTT AAT GCA GGG GAA TCA CGA CTC GCT GAA CTG TAC Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr 1780 1785 1790	5376

-120-

GTG GAA AGG GCA AAG TGG CTC TGG TCC AAG GGT GAT GTT CAC CAG GCA Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1795 1800 1805	5424
CTA ATT GTT CTT CAA AAA GGT GTT GAA TTA TGT TTT CCT GAA AAT GAA Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1810 1815 1820	5472
ACC CCA CCT GAG GGT AAG AAC ATG TTA ATC CAT GGT CGA GCT ATG CTA Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu 1825 1830 1835 1840	5520
CTA GTG GGC CGA TTT ATG GAA GAA ACA GCT AAC TTT GAA AGC AAT GCA Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1845 1850 1855	5568
ATT ATG AAA AAA TAT AAG GAT GTG ACC GCG TGC CTG CCA GAA TGG GAG Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1860 1865 1870	5616
GAT GGG CAT TTT TAC CTT GCC AAG TAC TAT GAC AAA TTG ATG CCC ATG Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1875 1880 1885	5664
GTC ACA GAC AAC AAA ATG GAA AAG CAA GGT GAT CTC ATC CGG TAT ATA Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1890 1895 1900	5712
GTT CTT CAT TTT GGC AGA TCT CTA CAA TAT GGA AAT CAG TTC ATA TAT Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1905 1910 1915 1920	5760
CAG TCA ATG CCA CGA ATG TTA ACT CTA TGG CTT GAT TAT GGT ACA AAG Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 1935	5808
GCA TAT GAA TGG GAA AAA GCT GGC CGC TCC GAT CGT GTA CAA ATG AGG Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1940 1945 1950	5856
AAT GAT TTG GGT AAA ATA AAC AAG GTT ATC ACA GAG CAT ACA AAC TAT Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 1965	5904
TTA GCT CCA TAT CAA TTT TTG ACT GCT TTT TCA CAA TTG ATC TCT CGA Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980	5952
ATT TGT CAT TCT CAC GAT GAA GTT TTT GTT GTG CTT GAT GGA AAT AAT Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn 1985 1990 1995 2000	6000
AGC CAA GTA TTT CTA GCC TAT CCT CAA CAA GCA ATG TGG ATG ATG ACA Ser Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr 2005 2010 2015	6048
GCT GTG TCA AAG TCA TCT TAT CCC ATG CGT GTG AAC AGA TGC AAG GAA Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030	6096
ATC CTC AAT AAA GCT ATT CAT ATG AAA AAA TCC TTA GAG AAG TTT GTT Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045	6144
GGA GAT GCA ACT CGC CTA ACA GAT AAG CTT CTA GAA TTG TGC AAT AAA Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys 2050 2055 2060	6192

-121-

CCG GTT GAT GGA AGT AGT TCC ACA TTA AGC ATG AGC ACT CAT TTT AAA	6240
Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys	
2065 2070 2075 2080	
ATG CTT AAA AAG CTG GTA GAA GAA GCA ACA TTT AGT GAA ATC CTC ATT	6288
Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile	
2085 2090 2095	
CCT CTA CAA TCA GTC ATG ATA CCT ACA CTT CCA TCA ATT CTG GGT ACC	6336
Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr	
2100 2105 2110	
CAT GCT AAC CAT GCT AGC CAT GAA CCA TTT CCT GGA CAT TGG GCC TAT	6384
His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr	
2115 2120 2125	
ATT GCA GGG TTT GAT GAT ATG GTG GAA ATT CTT GCT TCT CTT CAG AAA	6432
Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys	
2130 2135 2140	
CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC TAC ATC ATG	6480
Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met	
2145 2150 2155 2160	
ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA CTA ATG GAA	6528
Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu	
2165 2170 2175	
TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG TCT CGT	6576
Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg	
2180 2185 2190	
AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT	6624
Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp	
2195 2200 2205	
GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG AGA CCT	6672
Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro	
2210 2215 2220	
ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA	6720
Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys	
2225 2230 2235 2240	
GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA	6768
Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys	
2245 2250 2255	
CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT	6816
Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe	
2260 2265 2270	
CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT	6864
His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser	
2275 2280 2285	
AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT	6912
Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly	
2290 2295 2300	
TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT	6960
Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp	
2305 2310 2315 2320	
TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT	7008
Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn	
2325 2330 2335	

-122-

AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr 2340 2345 2350	7056
CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe 2355 2360 2365	7104
CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT CAG CGA GAG Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu 2370 2375 2380	7152
CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT CTT GTG GAA Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu 2385 2390 2395 2400	7200
TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG AAT GAA ACT Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr 2405 2410 2415	7248
GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT GAC ATT GAG Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu 2420 2425 2430	7296
CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG ACA GGA CTG Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu 2435 2440 2445	7344
CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA GAA GCT ACT Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr 2450 2455 2460	7392
GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT CCA TAT ATG Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met 2465 2470 2475 2480	7440
TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA AGTAAAAAAA AAAAAAAAAA	7500
AA	7502

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2480 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser
 1              5              10              15
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu
 20              25              30
Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val
 35              40              45
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu
 50              55              60
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile
 65              70              75              80

```


-123-

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly
 85 90 95
 Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu
 100 105 110
 Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu
 115 120 125
 Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu
 130 135 140
 Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys
 145 150 155 160
 Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp
 165 170 175
 Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe
 180 185 190
 Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val
 195 200 205
 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp
 210 215 220
 Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser
 225 230 235 240
 Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu
 245 250 255
 Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350
 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg
 370 375 380
 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp
 385 390 395 400
 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser
 405 410 415
 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu
 420 425 430

-124-

Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr
 435 440 445
 Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr
 450 455 460
 Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro
 465 470 475 480
 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu
 485 490 495
 Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe
 500 505 510
 Ile Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu
 515 520 525
 Ile Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala
 530 535 540
 Ser Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu
 545 550 555 560
 Thr Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu
 565 570 575
 Phe Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser
 580 585 590
 Gln Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys
 595 600 605
 Lys Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His
 610 615 620
 Leu Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala
 625 630 635 640
 Val Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val
 645 650 655
 Arg Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp
 660 665 670
 Ser Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu
 675 680 685
 Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr
 690 695 700
 Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu
 705 710 715 720
 Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser
 725 730 735
 Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala
 740 745 750
 Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro
 755 760 765
 Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala
 770 775 780

-125-

Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val
 785 790 795 800
 Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn
 805 810 815
 Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln
 820 825 830
 Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala
 835 840 845
 Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile
 850 855 860
 Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys
 865 870 875 880
 Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr
 885 890 895
 Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His
 900 905 910
 Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn
 915 920 925
 Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln
 930 935 940
 Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln
 945 950 955 960
 Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser
 965 970 975
 Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met
 980 985 990
 Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val
 995 1000 1005
 Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe
 1010 1015 1020
 Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp
 1025 1030 1035 1040
 His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu
 1045 1050 1055
 Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr
 1060 1065 1070
 Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile
 1075 1080 1085
 Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu
 1090 1095 1100
 Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr
 1105 1110 1115 1120
 Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg
 1125 1130 1135

-126-

Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu
 1140 1145 1150
 Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile
 1155 1160 1165
 Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser
 1170 1175 1180
 Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile
 1185 1190 1195 1200
 Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp
 1205 1210 1215
 Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu
 1220 1225 1230
 Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg
 1235 1240 1245
 Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr
 1250 1255 1260
 Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg
 1265 1270 1275 1280
 Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr
 1285 1290 1295
 Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys
 1300 1305 1310
 Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala
 1315 1320 1325
 Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser
 1330 1335 1340
 Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val
 1345 1350 1355 1360
 Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys
 1365 1370 1375
 Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu
 1380 1385 1390
 Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp
 1395 1400 1405
 Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu
 1410 1415 1420
 Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys
 1425 1430 1435 1440
 Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr
 1445 1450 1455
 Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile
 1460 1465 1470
 Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr
 1475 1480 1485

-127-

Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn
 1490 1495 1500
 Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His
 1505 1510 1515 1520
 Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro
 1525 1530 1535
 Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg
 1540 1545 1550
 Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln
 1555 1560 1565
 Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln
 1570 1575 1580
 Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser
 1585 1590 1595 1600
 Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys
 1605 1610 1615
 Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys
 1620 1625 1630
 Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys
 1635 1640 1645
 Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala
 1650 1655 1660
 Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr
 1665 1670 1675 1680
 Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu
 1685 1690 1695
 Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser
 1700 1705 1710
 Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn
 1715 1720 1725
 Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu
 1730 1735 1740
 Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp
 1745 1750 1755 1760
 Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala
 1765 1770 1775
 Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr
 1780 1785 1790
 Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala
 1795 1800 1805
 Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu
 1810 1815 1820
 Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu
 1825 1830 1835 1840

-128-

Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala
 1845 1850 1855
 Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu
 1860 1865 1870
 Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met
 1875 1880 1885
 Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile
 1890 1895 1900
 Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr
 1905 1910 1915 1920
 Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys
 1925 1930 1935
 Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg
 1940 1945 1950
 Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr
 1955 1960 1965
 Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg
 1970 1975 1980
 Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn
 1985 1990 1995 2000
 Ser Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr
 2005 2010 2015
 Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu
 2020 2025 2030
 Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val
 2035 2040 2045
 Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys
 2050 2055 2060
 Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys
 2065 2070 2075 2080
 Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile
 2085 2090 2095
 Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr
 2100 2105 2110
 His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr
 2115 2120 2125
 Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys
 2130 2135 2140
 Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met
 2145 2150 2155 2160
 Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu
 2165 2170 2175
 Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg
 2180 2185 2190

-129-

Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp
 2195 2200 2205
 Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro
 2210 2215 2220
 Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys
 2225 2230 2235 2240
 Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys
 2245 2250 2255
 Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe
 2260 2265 2270
 His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser
 2275 2280 2285
 Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly
 2290 2295 2300
 Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp
 2305 2310 2315 2320
 Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn
 2325 2330 2335
 Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr
 2340 2345 2350
 His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe
 2355 2360 2365
 Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu
 2370 2375 2380
 Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu
 2385 2390 2395 2400
 Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr
 2405 2410 2415
 Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu
 2420 2425 2430
 Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu
 2435 2440 2445
 Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr
 2450 2455 2460
 Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2465 2470 2475 2480

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 878 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

-130-

ATCCATTGTG TTGGAAAGGA ATGATGAATG TGGGATTATT GAATGGGTGA ACAATACTGC	60
TGGCTTGAGA CCTATTCTGA CCAAAATATA TAAAGAAAAG GGAGTGTATA TGACAGGAAA	120
GGAGCTTCGC CAGTGTATGC TACCAAAGTC AGCAGCTTTA TCTGAAAAAC TCAAAGTATT	180
CCAAGAATTA CTCCTGCCCA GGCATCCTCC TGTTTTTCAT GAGTGGTTTC TGAGAACATT	240
CCCTGATCCT ACATCATGGT ACAGTAGCAG ATCTGCATAT TGCCGCTCTA CTGCAGTCAT	300
GTCAATGGTT GGCTACATCC TGGGGCTTGG AGACCGTCAT GGTGAAAACA TTCTTTTGA	360
CTCTTTCCT GGTGAATGTG TACATGTAGA TTTCAACTGT CTMTTAAATA AGGGAGAAAC	420
GTTTGAAGTT CCGGAAATTG TACCATTTCG ACTGACTCAT AATATGGTTA ATGGAATGGG	480
TCCTATGGGA ACAGAGGGTC TATTTTCAAG AGCATGTGAA GTTACACTGA GACTGATGAG	540
GGATCAGAGA GAACCTTTAA TGAGTGTCTT AAAGACTTTT CTACACGATC CTCTAGTGGA	600
GTGGAGTAAA CCAGTGAAAG GACACTCCAA AGCACCCTG AATGAAACCG GGAAGTTGT	660
CAATGAGAAG GCCAAGACCC ATGTTCTTGA CATTGAACAA CGACTACAAG GTGTGATCAA	720
AACCCGAAAT AGAGTAACAG GGCTGCCATT ATCTATTGAA GGACATGTGC ATTACCTCAT	780
ACAAGAAGCT ACTGATGAAA ACTTACTCTG TCAGATGTAC CTTGGTTGGA CCCCATATAT	840
GTAAAATAAA ATTATTTCAA AGAAAAAAA AAAAAAAA	878

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7935 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
- (A) NAME/KEY: CDS
 - (B) LOCATION: 1..7932

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATG GGG GAA CAT GGC CTG GAG CTG GCT TCC ATG ATC CCC GCC CTG CGG	48
Met Gly Glu His Gly Leu Glu Leu Ala Ser Met Ile Pro Ala Leu Arg	
1 5 10 15	
GAG CTG GGC AGT GCC ACA CCA GAG GAA TAT AAT ACA GTT GTA CAG AAG	96
Glu Leu Gly Ser Ala Thr Pro Glu Glu Tyr Asn Thr Val Val Gln Lys	
20 25 30	

-131-

CCA AGA CAA ATT CTG TGT CAA TTC ATT GAC CGG ATA CTT ACA GAT GTA	144
Pro Arg Gln Ile Leu Cys Gln Phe Ile Asp Arg Ile Leu Thr Asp Val	
35 40 45	
AAT GTT GTT GCT GTA GAA CTT GTA AAG AAA ACT GAC TCT CAG CCA ACC	192
Asn Val Val Ala Val Glu Leu Val Lys Lys Thr Asp Ser Gln Pro Thr	
50 55 60	
TCC GTG ATG TTG CTT GAT TTC ATC CAG CAT ATC ATG AAA TCC TCC CCA	240
Ser Val Met Leu Leu Asp Phe Ile Gln His Ile Met Lys Ser Ser Pro	
65 70 75 80	
CTT ATG TTT GTA AAT GTG AGT GGA AGC CAT GAG CGC AAA GGC AGT TGT	288
Leu Met Phe Val Asn Val Ser Gly Ser His Glu Arg Lys Gly Ser Cys	
85 90 95	
ATT GAA TTC AGT AAT TGG ATC ATA ACG AGA CTT CTG CGG ATT GCA GCA	336
Ile Glu Phe Ser Asn Trp Ile Ile Thr Arg Leu Leu Arg Ile Ala Ala	
100 105 110	
ACT CCC TCC TGT CAT TTG TTA CAC AAG AAA ATC TGT GAA GTC ATC TGT	384
Thr Pro Ser Cys His Leu Leu His Lys Lys Ile Cys Glu Val Ile Cys	
115 120 125	
TCA TTA TTA TTT CTT TTT AAA AGC AAG AGT CCT GCT ATT TTT GGG GTA	432
Ser Leu Leu Phe Leu Phe Lys Ser Lys Ser Pro Ala Ile Phe Gly Val	
130 135 140	
CTC ACA AAA GAA TTA TTA CAA CTT TTT GAA GAC TTG GTT TAC CTC CAT	480
Leu Thr Lys Glu Leu Leu Gln Leu Phe Glu Asp Leu Val Tyr Leu His	
145 150 155 160	
AGA AGA AAT GTG ATG GGT CAT GCT GTG GAA TGG CCA GTG GTC ATG AGC	528
Arg Arg Asn Val Met Gly His Ala Val Glu Trp Pro Val Val Met Ser	
165 170 175	
CGA TTT TTA AGT CAA TTA GAT GAA CAC ATG GGA TAT TTA CAA TCA GCT	576
Arg Phe Leu Ser Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala	
180 185 190	
CCT TTG CAG TTG ATG AGT ATG CAA AAT TTA GAA TTT ATT GAA GTC ACT	624
Pro Leu Gln Leu Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr	
195 200 205	
TTA TTA ATG GTT CTT ACT CGT ATT ATT GCA ATT GTG TTT TTT AGA AGG	672
Leu Leu Met Val Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg	
210 215 220	
CAA GAA CTC TTA CTT TGG CAG ATA GGT TGT GTT CTG CTA GAG TAT GGT	720
Gln Glu Leu Leu Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly	
225 230 235 240	
AGT CCA AAA ATT AAA TCC CTA GCA ATT AGC TTT TTA ACA GAA CTT TTT	768
Ser Pro Lys Ile Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe	
245 250 255	
CAG CTT GGA GGA CTA CCA GCA CAA CCA GCT AGC ACT TTT TTC AGC TCA	816
Gln Leu Gly Gly Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser	
260 265 270	
TTT TTG GAA TTA TTA AAA CAC CTT GTA GAA ATG GAT ACT GAC CAA TTG	864
Phe Leu Glu Leu Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu	
275 280 285	
AAA CTC TAT GAA GAG CCA TTA TCA AAG CTG ATA AAG ACA CTA TTT CCC	912
Lys Leu Tyr Glu Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro	
290 295 300	

-132-

TTT GAA GCA GAA GCT TAT AGA AAT ATT GAA CCT GTC TAT TTA AAT ATG Phe Glu Ala Glu Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met 305 310 315 320	960
CTG CTG GAA AAA CTC TGT GTC ATG TTT GAA GAC GGT GTG CTC ATG CGG Leu Leu Glu Lys Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg 325 330 335	1008
CTT AAG TCT GAT TTG CTA AAA GCA GCT TTG TGC CAT TTA CTG CAG TAT Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr 340 345 350	1056
TTC CTT AAA TTT GTG CCA GCT GGG TAT GAA TCT GCT TTA CAA GTC AGG Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg 355 360 365	1104
AAG GTC TAT GTG AGA AAT ATT TGT AAA GCT CTT TTG GAT GTG CTT GGA Lys Val Tyr Val Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly 370 375 380	1152
ATT GAG GTA GAT GCA GAG TAC TTG TTG GGC CCA CTT TAT GCA GCT TTG Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu 385 390 395 400	1200
AAA ATG GAA AGT ATG GAA ATC ATT GAG GAG ATT CAA TGC CAA ACT CAA Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln 405 410 415	1248
CAG GAA AAC CTC AGC AGT AAT AGT GAT GGA ATA TCA CCC AAA AGG CGT Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg 420 425 430	1296
CGT CTC AGC TCG TCT CTA AAC CCT TCT AAA AGA GCA CCA AAA CAG ACT Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr 435 440 445	1344
GAG GAA ATT AAA CAT GTG GAC ATG AAC CAA AAG AGC ATA TTA TGG AGT Glu Glu Ile Lys His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser 450 455 460	1392
GCA CTG AAA CAG AAA GCT GAA TCC CTT CAG ATT TCC CTT GAA TAC AGT Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser 465 470 475 480	1440
GGC CTA AAG AAT CCT GTT ATT GAG ATG TTA GAA GGA ATT GCT GTT GTC Gly Leu Lys Asn Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val 485 490 495	1488
TTA CAA CTG ACT GCT CTG TGT ACT GTT CAT TGT TCT CAT CAA AAC ATG Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met 500 505 510	1536
AAC TGC CGT ACT TTC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAA CCT Asn Cys Arg Thr Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro 515 520 525	1584
TCT GTA GTG ATA ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT Ser Val Val Ile Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu 530 535 540	1632
AAG AGC TGT AGA AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG Lys Ser Cys Arg Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu 545 550 555 560	1680
GCA ACC ATT GAT AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG Ala Thr Ile Asp Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met 565 570 575	1728

-133-

CAA	GTA	AAC	AGT	TCA	TTT	GAA	GAT	CAT	ATC	CTG	GAA	GAT	TTA	TGT	GGT	1776
Gln	Val	Asn	Ser	Ser	Phe	Glu	Asp	His	Ile	Leu	Glu	Asp	Leu	Cys	Gly	
			580					585					590			
ATG	CTC	TCA	CTT	CCA	TGG	ATT	TAT	TCC	CAT	TCT	GAT	GAT	GGC	TGT	TTA	1824
Met	Leu	Ser	Leu	Pro	Trp	Ile	Tyr	Ser	His	Ser	Asp	Asp	Gly	Cys	Leu	
		595				600						605				
AAG	TTG	ACC	ACA	TTT	GCC	GCT	AAT	CTT	CTA	ACA	TTA	AGC	TGT	AGG	ATT	1872
Lys	Leu	Thr	Thr	Phe	Ala	Ala	Asn	Leu	Leu	Thr	Leu	Ser	Cys	Arg	Ile	
	610					615					620					
TCA	GAT	AGC	TAT	TCA	CCA	CAG	GCA	CAA	TCA	CGA	TGT	GTG	TTT	CTT	CTG	1920
Ser	Asp	Ser	Tyr	Ser	Pro	Gln	Ala	Gln	Ser	Arg	Cys	Val	Phe	Leu	Leu	
625					630					635					640	
ACT	CTG	TTT	CCA	AGA	AGA	ATA	TTC	CTT	GAG	TGG	AGA	ACA	GCA	GTT	TAC	1968
Thr	Leu	Phe	Pro	Arg	Arg	Ile	Phe	Leu	Glu	Trp	Arg	Thr	Ala	Val	Tyr	
				645					650					655		
AAC	TGG	GCC	CTG	CAG	AGC	TCC	CAT	GAA	GTA	ATC	CGG	GCT	AGT	TGT	GTT	2016
Asn	Trp	Ala	Leu	Gln	Ser	Ser	His	Glu	Val	Ile	Arg	Ala	Ser	Cys	Val	
		660						665					670			
AGT	GGA	TTT	TTT	ATC	TTA	TTG	CAG	CAG	CAG	AAT	TCT	TGT	AAC	AGA	GTT	2064
Ser	Gly	Phe	Phe	Ile	Leu	Leu	Gln	Gln	Gln	Asn	Ser	Cys	Asn	Arg	Val	
		675					680					685				
CCC	AAG	ATT	CTT	ATA	GAT	AAA	GTC	AAA	GAT	GAT	TCT	GAC	ATT	GTC	AAG	2112
Pro	Lys	Ile	Leu	Ile	Asp	Lys	Val	Lys	Asp	Asp	Ser	Asp	Ile	Val	Lys	
	690					695					700					
AAA	GAA	TTT	GCT	TCT	ATA	CTT	GGT	CAA	CTT	GTC	TGT	ACT	CTT	CAC	GGC	2160
Lys	Glu	Phe	Ala	Ser	Ile	Leu	Gly	Gln	Leu	Val	Cys	Thr	Leu	His	Gly	
705					710					715					720	
ATG	TTT	TAT	CTG	ACA	AGT	TCT	TTA	ACA	GAA	CCT	TTC	TCT	GAA	CAC	GGA	2208
Met	Phe	Tyr	Leu	Thr	Ser	Ser	Leu	Thr	Glu	Pro	Phe	Ser	Glu	His	Gly	
				725					730					735		
CAT	GTG	GAC	CTC	TTC	TGT	AGG	AAC	TTG	AAA	GCC	ACT	TCT	CAA	CAT	GAA	2256
His	Val	Asp	Leu	Phe	Cys	Arg	Asn	Leu	Lys	Ala	Thr	Ser	Gln	His	Glu	
			740					745					750			
TGT	TCA	TCT	TCT	CAA	CTA	AAA	GCT	TCT	GTC	TGC	AAG	CCA	TTC	CTT	TTC	2304
Cys	Ser	Ser	Ser	Gln	Leu	Lys	Ala	Ser	Val	Cys	Lys	Pro	Phe	Leu	Phe	
		755					760					765				
CTA	CTG	AAA	AAA	AAA	ATA	CCT	AGT	CCA	GTA	AAA	CTT	GCT	TTC	ATA	GAT	2352
Leu	Leu	Lys	Lys	Lys	Ile	Pro	Ser	Pro	Val	Lys	Leu	Ala	Phe	Ile	Asp	
	770					775					780					
AAT	CTA	CAT	CAT	CTT	TGT	AAG	CAT	CTT	GAT	TTT	AGA	GAA	GAT	GAA	ACA	2400
Asn	Leu	His	His	Leu	Cys	Lys	His	Leu	Asp	Phe	Arg	Glu	Asp	Glu	Thr	
785					790					795					800	
GAT	GTA	AAA	GCA	GTT	CTT	GGA	ACT	TTA	TTA	AAT	TTA	ATG	GAA	GAT	CCA	2448
Asp	Val	Lys	Ala	Val	Leu	Gly	Thr	Leu	Leu	Asn	Leu	Met	Glu	Asp	Pro	
			805						810					815		
GAC	AAA	GAT	GTT	AGA	GTG	GCT	TTT	AGT	GGA	AAT	ATC	AAG	CAC	ATA	TTG	2496
Asp	Lys	Asp	Val	Arg	Val	Ala	Phe	Ser	Gly	Asn	Ile	Lys	His	Ile	Leu	
			820					825					830			
GAA	TCC	TTG	GAC	TCT	GAA	GAT	GGA	TTT	ATA	AAG	GAG	CTT	TTT	GTC	TTA	2544
Glu	Ser	Leu	Asp	Ser	Glu	Asp	Gly	Phe	Ile	Lys	Glu	Leu	Phe	Val	Leu	
		835					840					845				

-134-

AGA ATG AAG GAA GCA TAT ACA CAT GCC CAA ATA TCA AGA AAT AAT GAG Arg Met Lys Glu Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu 850 855 860	2592
CTG AAG GAT ACC TTG ATT CTT ACA ACA GGG GAT ATT GGA AGG GCC GCA Leu Lys Asp Thr Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala 865 870 875 880	2640
AAA GGA GAT TTG GTA CCA TTT GCA CTC TTA CAC TTA TTG CAT TGT TTG Lys Gly Asp Leu Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu 885 890 895	2688
TTA TCC AAG TCA GCA TCT GTC TCT GGA GCA GCA TAC ACA GAA ATT AGA Leu Ser Lys Ser Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg 900 905 910	2736
GCT CTG GTT GCA GCT AAA AGT GTT AAA CTG CAA AGT TTT TTC AGC CAG Ala Leu Val Ala Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln 915 920 925	2784
TAT AAG AAA CCC ATC TGT CAG TTT TTG GTA GAA TCC CTT CAC TCT AGT Tyr Lys Lys Pro Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser 930 935 940	2832
CAG ATG ACA GCA CTT CCG AAT ACT CCA TGC CAG AAT GCT GAC GTG CGA Gln Met Thr Ala Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg 945 950 955 960	2880
AAA CAA GAT GTG GCT CAC CAG AGA GAA ATG GCT TTA AAT ACG TTG TCT Lys Gln Asp Val Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser 965 970 975	2928
GAA ATT GCC AAC GTT TTC GAC TTT CCT GAT CTT AAT CGT TTT CTT ACT Glu Ile Ala Asn Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr 980 985 990	2976
AGG ACA TTA CAA GTT CTA CTA CCT GAT CTT GCT GCC AAA GCA AGC CCT Arg Thr Leu Gln Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro 995 1000 1005	3024
GCA GCT TCT GCT CTC ATT CGA ACT TTA GGA AAA CAA TTA AAT GTC AAT Ala Ala Ser Ala Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn 1010 1015 1020	3072
CGT AGA GAG ATT TTA ATA AAC AAC TTC AAA TAT ATT TTT TCT CAT TTG Arg Arg Glu Ile Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu 1025 1030 1035 1040	3120
GTC TGT TCT TGT TCC AAA GAT GAA TTA GAA CGT GCC CTT CAT TAT CTG Val Cys Ser Cys Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu 1045 1050 1055	3168
AAG AAT GAA ACA GAA ATT GAA CTG GGG AGC CTG TTG AGA CAA GAT TTC Lys Asn Glu Thr Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe 1060 1065 1070	3216
CAA GGA TTG CAT AAT GAA TTA TTG CTG CGT ATT GGA GAA CAC TAT CAA Gln Gly Leu His Asn Glu Leu Leu Arg Ile Gly Glu His Tyr Gln 1075 1080 1085	3264
CAG GTT TTT AAT GGT TTG TCA ATA CTT GCC TCA TTT GCA TCC AGT GAT Gln Val Phe Asn Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp 1090 1095 1100	3312
GAT CCA TAT CAG GGC CCG AGA GAT ATC ATA TCA CCT GAA CTG ATG GCT Asp Pro Tyr Gln Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala 1105 1110 1115 1120	3360

-135-

GAT TAT TTA CAA CCC AAA TTG TTG GGC ATT TTG GCT TTT TTT AAC ATG Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met 1125 1130 1135	3408
CAG TTA CTG AGC TCT AGT GTT GGC ATT GAA GAT AAG AAA ATG GCC TTG Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu 1140 1145 1150	3456
AAC AGT TTG ATG TCT TTG ATG AAG TTA ATG GGA CCC AAA CAT GTC AGT Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser 1155 1160 1165	3504
TCT GTG AGG GTG AAG ATG ATG ACC ACA CTG AGA ACT GGC CTT CGA TTC Ser Val Arg Val Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe 1170 1175 1180	3552
AAG GAT GAT TTT CCT GAA TTG TGT TGC AGA GCT TGG GAC TGC TTT GTT Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val 1185 1190 1195 1200	3600
CGC TGC CTG GAT CAT GCT TGT CTG GGC TCC CTT CTC AGT CAT GTA ATA Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile 1205 1210 1215	3648
GTA GCT TTG TTA CCT CTT ATA CAC ATC CAG CCT AAA GAA ACT GCA GCT Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala 1220 1225 1230	3696
ATC TTC CAC TAC CTC ATA ATT GAA AAC AGG GAT GCT GTG CAA GAT TTT Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe 1235 1240 1245	3744
CTT CAT GAA ATA TAT TTT TTA CCT GAT CAT CCA GAA TTA AAA AAG ATA Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile 1250 1255 1260	3792
AAA GCC GTT CTC CAG GAA TAC AGA AAG GAG ACC TCT GAG AGC ACT GAT Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp 1265 1270 1275 1280	3840
CTT CAG ACA ACT CTT CAG CTC TCT ATG AAG GCC ATT CAA CAT GAA AAT Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn 1285 1290 1295	3888
GTC GAT GTT CGT ATT CAT GCT CTT ACA AGC TTG AAG GAA ACC TTG TAT Val Asp Val Arg Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr 1300 1305 1310	3936
AAA AAT CAG GAA AAA CTG ATA AAG TAT GCA ACA GAC AGT GAA ACA GTA Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val 1315 1320 1325	3984
GAA CCT ATT ATC TCA CAG TTG GTG ACA GTG CTT TTG AAA GGT TGC CAA Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln 1330 1335 1340	4032
GAT GCA AAC TCT CAA GCT CGG TTG CTC TGT GGG GAA TGT TTA GGG GAA Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu 1345 1350 1355 1360	4080
TTG GGG GCG ATA GAT CCA GGT CGA TTA GAT TTC TCA ACA ACT GAA ACT Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr 1365 1370 1375	4128
CAA GGA AAA GAT TTT ACA TTT GTG ACT GGA GTA GAA GAT TCA AGC TTT Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe 1380 1385 1390	4176

-136-

GCC TAT GGA TTA TTG ATG GAG CTA ACA AGA GCT TAC CTT GCG TAT GCT Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala 1395 1400 1405	4224
GAT AAT AGC CGA GCT CAA GAT TCA GCT GCC TAT GCC ATT CAG GAG TTG Asp Asn Ser Arg Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu 1410 1415 1420	4272
CTT TCT ATT TAT GAC TGT AGA GAG ATG GAG ACC AAC GGC CCA GGT CAC Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His 1425 1430 1435 1440	4320
CAA TTG TGG AGG AGA TTT CCT GAG CAT GTT CGG GAA ATA CTA GAA CCT Gln Leu Trp Arg Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro 1445 1450 1455	4368
CAT CTA AAT ACC AGA TAC AAG AGT TCT CAG AAG TCA ACC GAT TGG TCT His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser 1460 1465 1470	4416
GGA GTA AAG AAG CCA ATT TAC TTA AGT AAA TTG GGT AGT AAC TTT GCA Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala 1475 1480 1485	4464
GAA TGG TCA GCA TCT TGG GCA GGT TAT CTT ATT ACA AAG GTT CGA CAT Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His 1490 1495 1500	4512
GAT CTT GCC AGT AAA ATT TTC ACC TGC TGT AGC ATT ATG ATG AAG CAT Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His 1505 1510 1515 1520	4560
GAT TTC AAA GTG ACC ATC TAT CTT CTT CCA CAT ATT CTG GTG TAT GTC Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val 1525 1530 1535	4608
TTA CTG GGT TGT AAT CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile 1540 1545 1550	4656
ATG GCA GTT CTA AAG CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC Met Ala Val Leu Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp 1555 1560 1565	4704
ATT GCA TCT GAT CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met 1570 1575 1580	4752
CTT GAC CAT CTC ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA Leu Asp His Leu Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys 1585 1590 1595 1600	4800
GCT GAG AAA TGT CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser 1605 1610 1615	4848
ATG GTA TCT ACT GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe 1620 1625 1630	4896
CTA GAC CTC ATA CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser 1635 1640 1645	4944
AAA GCA TAC ACA CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu 1650 1655 1660	4992

-137-

AAG Lys 1665	AAG Lys 1665	CAA Gln 1665	AAT Asn 1665	ATT Ile 1665	CAG Gln 1670	GAA Glu 1670	CAT His 1670	CTT Leu 1670	GGA Gly 1675	TTT Phe 1675	TTA Leu 1675	CAG Gln 1675	AAA Lys 1680	TTG Leu 1680	TAT Tyr 1680	5040
GCT Ala 1685	GCT Ala 1685	ATG Met 1685	CAT His 1685	GAA Glu 1685	CCT Pro 1685	GAT Asp 1685	GGA Gly 1690	GTG Val 1690	GCC Ala 1690	GGA Gly 1695	GTC Val 1695	AGT Ser 1695	GCA Ala 1695	ATT Ile 1695	AGA Arg 1695	5088
AAG Lys 1700	GCA Ala 1700	GAA Glu 1700	CCA Pro 1700	TCT Ser 1700	CTA Leu 1700	AAA Lys 1705	GAA Glu 1705	CAG Gln 1705	ATC Ile 1705	CTT Leu 1710	GAA Glu 1710	CAT His 1710	GAA Glu 1710	AGC Ser 1710	CTT Leu 1710	5136
GGC Gly 1715	TTG Leu 1715	CTG Leu 1715	AGG Arg 1715	GAT Asp 1715	GCC Ala 1720	ACT Thr 1720	GCT Ala 1720	TGT Cys 1720	TAT Tyr 1725	GAC Asp 1725	AGG Arg 1725	GCT Ala 1725	ATT Ile 1725	CAG Gln 1725	CTA Leu 1725	5184
GAA Glu 1730	CCA Pro 1730	GAC Asp 1730	CAG Gln 1735	ATC Ile 1735	ATT Ile 1735	CAT His 1735	TAT Tyr 1735	CAT His 1740	GGT Gly 1740	GTA Val 1740	GTA Val 1740	AAG Lys 1740	TCC Ser 1740	ATG Met 1740	TTA Leu 1740	5232
GGT Gly 1745	CTT Leu 1745	GGT Gly 1745	CAG Gln 1750	CTG Leu 1750	TCT Ser 1750	ACT Thr 1750	GTT Val 1750	ATC Ile 1755	ACT Thr 1755	CAG Gln 1755	GTG Val 1755	AAT Asn 1760	GGA Gly 1760	GTG Val 1760	CAT His 1760	5280
GCT Ala 1765	AAC Asn 1765	AGG Arg 1765	TCC Ser 1765	GAG Glu 1765	TGG Trp 1765	ACA Thr 1770	GAT Asp 1770	GAA Glu 1770	TTA Leu 1770	AAC Asn 1770	ACG Thr 1770	TAC Tyr 1775	AGA Arg 1775	GTG Val 1775	GAA Glu 1775	5328
GCA Ala 1780	GCT Ala 1780	TGG Trp 1780	AAA Lys 1780	TTG Leu 1780	TCA Ser 1785	CAG Gln 1785	TGG Trp 1785	GAT Asp 1785	TTG Leu 1785	GTG Val 1790	GAA Glu 1790	AAC Asn 1790	TAT Tyr 1790	TTG Leu 1790	GCA Ala 1790	5376
GCA Ala 1795	GAT Asp 1795	GGA Gly 1795	AAA Lys 1800	TCT Ser 1800	ACA Thr 1800	ACA Thr 1800	TGG Trp 1800	AGT Ser 1805	GTC Val 1805	AGA Arg 1805	CTG Leu 1805	GGA Gly 1805	CAG Gln 1805	CTA Leu 1805	TTA Leu 1805	5424
TTA Leu 1810	TCA Ser 1810	GCC Ala 1815	AAA Lys 1815	AAA Lys 1815	AGA Arg 1815	GAT Asp 1815	ATC Ile 1820	ACA Thr 1820	GCT Ala 1820	TTT Phe 1820	TAT Tyr 1820	GAC Asp 1820	TCA Ser 1820	CTG Leu 1820	AAA Lys 1820	5472
CTA Leu 1825	GTG Val 1825	AGA Arg 1830	GCA Ala 1830	GAA Glu 1830	CAA Gln 1830	ATT Ile 1835	GTA Val 1835	CCT Pro 1835	CTT Leu 1835	TCA Ser 1835	GCT Ala 1840	AGC Ala 1840	TTT Phe 1840	GAA Glu 1840		5520
AGA Arg 1845	GGC Gly 1845	TCC Ser 1845	TAC Tyr 1845	CAA Gln 1845	CGA Arg 1850	GGA Gly 1850	TAT Tyr 1850	GAA Glu 1850	TAT Tyr 1850	ATT Ile 1850	GTG Val 1855	AGA Arg 1855	TTG Leu 1855	CAC His 1855	ATG Met 1855	5568
TTA Leu 1860	TGT Cys 1860	GAG Glu 1860	TTG Leu 1860	GAG Glu 1860	CAT His 1865	AGC Ser 1865	ATC Ile 1865	AAA Lys 1865	CCA Pro 1865	CTT Leu 1870	TTC Phe 1870	CAG Gln 1870	CAT His 1870	TCT Ser 1870	CCA Pro 1870	5616
GGT Gly 1875	GAC Asp 1875	AGT Ser 1875	TCT Ser 1880	CAA Gln 1880	GAA Glu 1880	GAT Asp 1880	TCT Ser 1880	CTA Leu 1885	AAC Asn 1885	TGG Trp 1885	GTA Val 1885	GCT Ala 1885	CGA Arg 1885	CTA Leu 1885	GAA Glu 1885	5664
ATG Met 1890	ACC Thr 1890	CAG Gln 1895	AAT Asn 1895	TCC Ser 1895	TAC Tyr 1895	AGA Arg 1895	GCC Ala 1900	AAG Lys 1900	GAG Glu 1900	CCT Pro 1900	ATC Ile 1900	CTG Leu 1900	GCT Ala 1900	CTC Leu 1900	CGG Arg 1900	5712
AGG Arg 1905	GCT Ala 1905	TTA Leu 1910	CTA Leu 1910	AGC Ser 1910	CTC Leu 1910	AAC Asn 1915	AAA Lys 1915	AGA Arg 1915	CCA Pro 1915	GAT Asp 1915	TAC Tyr 1920	AAT Asn 1920	GAA Glu 1920	ATG Met 1920	GTT Val 1920	5760
GGA Gly 1925	GAA Glu 1925	TGC Cys 1925	TGG Trp 1925	CTG Leu 1925	CAG Gln 1930	AGT Ser 1930	GCC Ala 1930	AGG Arg 1930	GTA Val 1930	GCT Ala 1930	AGA Arg 1935	AAG Lys 1935	GCT Ala 1935	GGT Gly 1935	CAC His 1935	5808

-138-

CAC CAG ACA GCC TAC AAT GCT CTC CTT AAT GCA GGG GAA TCA CGA CTC His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu 1940 1945 1950	5856
GCT GAA CTG TAC GTG GAA AGG GCA AAG TGG CTC TGG TCC AAG GGT GAT Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp 1955 1960 1965	5904
GTT CAC CAG GCA CTA ATT GTT CTT CAA AAA GGT GTT GAA TTA TGT TTT Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe 1970 1975 1980	5952
CCT GAA AAT GAA ACC CCA CCT GAG GGT AAG AAC ATG TTA ATC CAT GGT Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly 1985 1990 1995 2000	6000
CGA GCT ATG CTA CTA GTG GGC CGA TTT ATG GAA GAA ACA GCT AAC TTT Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe 2005 2010 2015	6048
GAA AGC AAT GCA ATT ATG AAA AAA TAT AAG GAT GTG ACC GCG TGC CTG Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu 2020 2025 2030	6096
CCA GAA TGG GAG GAT GGG CAT TTT TAC CTT GCC AAG TAC TAT GAC AAA Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys 2035 2040 2045	6144
TTG ATG CCC ATG GTC ACA GAC AAC AAA ATG GAA AAG CAA GGT GAT CTC Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu 2050 2055 2060	6192
ATC CGG TAT ATA GTT CTT CAT TTT GGC AGA TCT CTA CAA TAT GGA AAT Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn 2065 2070 2075 2080	6240
CAG TTC ATA TAT CAG TCA ATG CCA CGA ATG TTA ACT CTA TGG CTT GAT Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp 2085 2090 2095	6288
TAT GGT ACA AAG GCA TAT GAA TGG GAA AAA GCT GGC CGC TCC GAT CGT Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg 2100 2105 2110	6336
GTA CAA ATG AGG AAT GAT TTG GGT AAA ATA AAC AAG GTT ATC ACA GAG Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu 2115 2120 2125	6384
CAT ACA AAC TAT TTA GCT CCA TAT CAA TTT TTG ACT GCT TTT TCA CAA His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln 2130 2135 2140	6432
TTG ATC TCT CGA ATT TGT CAT TCT CAC GAT GAA GTT TTT GTT GTC TTG Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val Phe Val Val Leu 2145 2150 2155 2160	6480
ATG GAA ATA ATA GCC AAA GTA TTT CTA GCC TAT CCT CAA CAA GCA ATG Met Glu Ile Ile Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met 2165 2170 2175	6528
TGG ATG ATG ACA GCT GTG TCA AAG TCA TCT TAT CCC ATG CGT GTG AAC Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn 2180 2185 2190	6576
AGA TGC AAG GAA ATC CTC AAT AAA GCT ATT CAT ATG AAA AAA TCC TTA Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu 2195 2200 2205	6624

-139-

GAG AAG TTT GTT GGA GAT GCA ACT CGC CTA ACA GAT AAG CTT CTA GAA Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu 2210 2215 2220	6672
TTG TGC AAT AAA CCG GTT GAT GGA AGT AGT TCC ACA TTA AGC ATG AGC Leu Cys Asn Lys Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser 2225 2230 2235 2240	6720
ACT CAT TTT AAA ATG CTT AAA AAG CTG GTA GAA GAA GCA ACA TTT AGT Thr His Phe Lys Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser 2245 2250 2255	6768
GAA ATC CTC ATT CCT CTA CAA TCA GTC ATG ATA CCT ACA CTT CCA TCA Glu Ile Leu Ile Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser 2260 2265 2270	6816
ATT CTG GGT ACC CAT GCT AAC CAT GCT AGC CAT GAA CCA TTT CCT GGA Ile Leu Gly Thr His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly 2275 2280 2285	6864
CAT TGG GCC TAT ATT GCA GGG TTT GAT GAT ATG GTG GAA ATT CTT GCT His Trp Ala Tyr Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala 2290 2295 2300	6912
TCT CTT CAG AAA CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys 2305 2310 2315 2320	6960
TTC TAC ATC ATG ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys 2325 2330 2335	7008
AGA CTA ATG GAA TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp 2340 2345 2350	7056
GCA GAG TCT CGT AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT Ala Glu Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile 2355 2360 2365	7104
CCA CTA AAT GAT GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala 2370 2375 2380	7152
GGT TTG AGA CCT ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr 2385 2390 2395 2400	7200
ATG ACA GGA AAA GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala 2405 2410 2415	7248
TTA TCT GAA AAA CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His 2420 2425 2430	7296
CCT CCT ATT TTT CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr 2435 2440 2445	7344
TCA TGG TAC AGT AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met 2450 2455 2460	7392
TCA ATG GTT GGT TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn 2465 2470 2475 2480	7440

-140-

ATT CTC TTT GAT TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn 2485 2490 2495	7488
TGT CTT TTC AAT AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro 2500 2505 2510	7536
TTT CGC CTG ACT CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA Phe Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr 2515 2520 2525	7584
GAG GGT CTT TTT CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg 2530 2535 2540	7632
GAT CAG CGA GAG CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp 2545 2550 2555 2560	7680
CCT CTT GTG GAA TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro 2565 2570 2575	7728
CTG AAT GAA ACT GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val 2580 2585 2590	7776
CTT GAC ATT GAG CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg 2595 2600 2605	7824
GTG ACA GGA CTG CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile 2610 2615 2620	7872
CAA GAA GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp 2625 2630 2635 2640	7920
ACT CCA TAT ATG TGA Thr Pro Tyr Met	7935

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2644 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gly Glu His Gly Leu Glu Leu Ala Ser Met Ile Pro Ala Leu Arg
1 5 10 15

Glu Leu Gly Ser Ala Thr Pro Glu Glu Tyr Asn Thr Val Val Gln Lys
20 25 30

Pro Arg Gln Ile Leu Cys Gln Phe Ile Asp Arg Ile Leu Thr Asp Val
35 40 45

-141-

Asn	Val	Val	Ala	Val	Glu	Leu	Val	Lys	Lys	Thr	Asp	Ser	Gln	Pro	Thr	50	55	60
Ser	Val	Met	Leu	Leu	Asp	Phe	Ile	Gln	His	Ile	Met	Lys	Ser	Ser	Pro	65	70	75
Leu	Met	Phe	Val	Asn	Val	Ser	Gly	Ser	His	Glu	Arg	Lys	Gly	Ser	Cys	85	90	95
Ile	Glu	Phe	Ser	Asn	Trp	Ile	Ile	Thr	Arg	Leu	Leu	Arg	Ile	Ala	Ala	100	105	110
Thr	Pro	Ser	Cys	His	Leu	Leu	His	Lys	Lys	Ile	Cys	Glu	Val	Ile	Cys	115	120	125
Ser	Leu	Leu	Phe	Leu	Phe	Lys	Ser	Lys	Ser	Pro	Ala	Ile	Phe	Gly	Val	130	135	140
Leu	Thr	Lys	Glu	Leu	Leu	Gln	Leu	Phe	Glu	Asp	Leu	Val	Tyr	Leu	His	145	150	155
Arg	Arg	Asn	Val	Met	Gly	His	Ala	Val	Glu	Trp	Pro	Val	Val	Met	Ser	165	170	175
Arg	Phe	Leu	Ser	Gln	Leu	Asp	Glu	His	Met	Gly	Tyr	Leu	Gln	Ser	Ala	180	185	190
Pro	Leu	Gln	Leu	Met	Ser	Met	Gln	Asn	Leu	Glu	Phe	Ile	Glu	Val	Thr	195	200	205
Leu	Leu	Met	Val	Leu	Thr	Arg	Ile	Ile	Ala	Ile	Val	Phe	Phe	Arg	Arg	210	215	220
Gln	Glu	Leu	Leu	Leu	Trp	Gln	Ile	Gly	Cys	Val	Leu	Leu	Glu	Tyr	Gly	225	230	235
Ser	Pro	Lys	Ile	Lys	Ser	Leu	Ala	Ile	Ser	Phe	Leu	Thr	Glu	Leu	Phe	245	250	255
Gln	Leu	Gly	Gly	Leu	Pro	Ala	Gln	Pro	Ala	Ser	Thr	Phe	Phe	Ser	Ser	260	265	270
Phe	Leu	Glu	Leu	Leu	Lys	His	Leu	Val	Glu	Met	Asp	Thr	Asp	Gln	Leu	275	280	285
Lys	Leu	Tyr	Glu	Glu	Pro	Leu	Ser	Lys	Leu	Ile	Lys	Thr	Leu	Phe	Pro	290	295	300
Phe	Glu	Ala	Glu	Ala	Tyr	Arg	Asn	Ile	Glu	Pro	Val	Tyr	Leu	Asn	Met	305	310	315
Leu	Leu	Glu	Lys	Leu	Cys	Val	Met	Phe	Glu	Asp	Gly	Val	Leu	Met	Arg	325	330	335
Leu	Lys	Ser	Asp	Leu	Leu	Lys	Ala	Ala	Leu	Cys	His	Leu	Leu	Gln	Tyr	340	345	350
Phe	Leu	Lys	Phe	Val	Pro	Ala	Gly	Tyr	Glu	Ser	Ala	Leu	Gln	Val	Arg	355	360	365
Lys	Val	Tyr	Val	Arg	Asn	Ile	Cys	Lys	Ala	Leu	Leu	Asp	Val	Leu	Gly	370	375	380
Ile	Glu	Val	Asp	Ala	Glu	Tyr	Leu	Leu	Gly	Pro	Leu	Tyr	Ala	Ala	Leu	385	390	395

-142-

Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln
 405 410 415
 Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg
 420 425 430
 Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr
 435 440 445
 Glu Glu Ile Lys His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser
 450 455 460
 Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser
 465 470 475 480
 Gly Leu Lys Asn Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val
 485 490 495
 Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met
 500 505 510
 Asn Cys Arg Thr Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro
 515 520 525
 Ser Val Val Ile Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu
 530 535 540
 Lys Ser Cys Arg Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu
 545 550 555 560
 Ala Thr Ile Asp Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met
 565 570 575
 Gln Val Asn Ser Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly
 580 585 590
 Met Leu Ser Leu Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu
 595 600 605
 Lys Leu Thr Thr Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile
 610 615 620
 Ser Asp Ser Tyr Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu
 625 630 635 640
 Thr Leu Phe Pro Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr
 645 650 655
 Asn Trp Ala Leu Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val
 660 665 670
 Ser Gly Phe Phe Ile Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val
 675 680 685
 Pro Lys Ile Leu Ile Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys
 690 695 700
 Lys Glu Phe Ala Ser Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly
 705 710 715 720
 Met Phe Tyr Leu Thr Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly
 725 730 735
 His Val Asp Leu Phe Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu
 740 745 750

-143-

Cys Ser Ser Ser Gln Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe
 755 760 765
 Leu Leu Lys Lys Lys Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp
 770 775 780
 Asn Leu His His Leu Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr
 785 790 795 800
 Asp Val Lys Ala Val Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro
 805 810 815
 Asp Lys Asp Val Arg Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu
 820 825 830
 Glu Ser Leu Asp Ser Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu
 835 840 845
 Arg Met Lys Glu Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu
 850 855 860
 Leu Lys Asp Thr Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala
 865 870 875 880
 Lys Gly Asp Leu Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu
 885 890 895
 Leu Ser Lys Ser Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg
 900 905 910
 Ala Leu Val Ala Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln
 915 920 925
 Tyr Lys Lys Pro Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser
 930 935 940
 Gln Met Thr Ala Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg
 945 950 955 960
 Lys Gln Asp Val Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser
 965 970 975
 Glu Ile Ala Asn Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr
 980 985 990
 Arg Thr Leu Gln Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro
 995 1000 1005
 Ala Ala Ser Ala Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn
 1010 1015 1020
 Arg Arg Glu Ile Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu
 1025 1030 1035 1040
 Val Cys Ser Cys Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu
 1045 1050 1055
 Lys Asn Glu Thr Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe
 1060 1065 1070
 Gln Gly Leu His Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln
 1075 1080 1085
 Gln Val Phe Asn Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp
 1090 1095 1100

-144-

Asp Pro Tyr Gln Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala
 1105 1110 1115 1120
 Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met
 1125 1130 1135
 Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu
 1140 1145 1150
 Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser
 1155 1160 1165
 Ser Val Arg Val Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe
 1170 1175 1180
 Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val
 1185 1190 1195 1200
 Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile
 1205 1210 1215
 Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala
 1220 1225 1230
 Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe
 1235 1240 1245
 Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile
 1250 1255 1260
 Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp
 1265 1270 1275 1280
 Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn
 1285 1290 1295
 Val Asp Val Arg Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr
 1300 1305 1310
 Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val
 1315 1320 1325
 Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln
 1330 1335 1340
 Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu
 1345 1350 1355 1360
 Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr
 1365 1370 1375
 Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe
 1380 1385 1390
 Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala
 1395 1400 1405
 Asp Asn Ser Arg Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu
 1410 1415 1420
 Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His
 1425 1430 1435 1440
 Gln Leu Trp Arg Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro
 1445 1450 1455

-145-

His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser
 1460 1465 1470
 Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala
 1475 1480 1485
 Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His
 1490 1495 1500
 Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His
 1505 1510 1515 1520
 Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val
 1525 1530 1535
 Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile
 1540 1545 1550
 Met Ala Val Leu Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp
 1555 1560 1565
 Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met
 1570 1575 1580
 Leu Asp His Leu Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys
 1585 1590 1595 1600
 Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser
 1605 1610 1615
 Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe
 1620 1625 1630
 Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser
 1635 1640 1645
 Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu
 1650 1655 1660
 Lys Lys Gln Asn Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr
 1665 1670 1675 1680
 Ala Ala Met His Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg
 1685 1690 1695
 Lys Ala Glu Pro Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu
 1700 1705 1710
 Gly Leu Leu Arg Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu
 1715 1720 1725
 Glu Pro Asp Gln Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu
 1730 1735 1740
 Gly Leu Gly Gln Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His
 1745 1750 1755 1760
 Ala Asn Arg Ser Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu
 1765 1770 1775
 Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala
 1780 1785 1790
 Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu
 1795 1800 1805

-146-

Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys
 1810 1815 1820
 Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu
 1825 1830 1835 1840
 Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met
 1845 1850 1855
 Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro
 1860 1865 1870
 Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu
 1875 1880 1885
 Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg
 1890 1895 1900
 Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val
 1905 1910 1915 1920
 Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His
 1925 1930 1935
 His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu
 1940 1945 1950
 Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp
 1955 1960 1965
 Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe
 1970 1975 1980
 Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly
 1985 1990 1995 2000
 Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe
 2005 2010 2015
 Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu
 2020 2025 2030
 Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys
 2035 2040 2045
 Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu
 2050 2055 2060
 Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn
 2065 2070 2075 2080
 Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp
 2085 2090 2095
 Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg
 2100 2105 2110
 Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu
 2115 2120 2125
 His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln
 2130 2135 2140
 Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val Phe Val Val Leu
 2145 2150 2155 2160

-147-

Met Glu Ile Ile Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met
 2165 2170 2175
 Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn
 2180 2185 2190
 Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu
 2195 2200 2205
 Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu
 2210 2215 2220
 Leu Cys Asn Lys Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser
 2225 2230 2235 2240
 Thr His Phe Lys Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser
 2245 2250 2255
 Glu Ile Leu Ile Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser
 2260 2265 2270
 Ile Leu Gly Thr His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly
 2275 2280 2285
 His Trp Ala Tyr Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala
 2290 2295 2300
 Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys
 2305 2310 2315 2320
 Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys
 2325 2330 2335
 Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp
 2340 2345 2350
 Ala Glu Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile
 2355 2360 2365
 Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala
 2370 2375 2380
 Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr
 2385 2390 2395 2400
 Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala
 2405 2410 2415
 Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His
 2420 2425 2430
 Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr
 2435 2440 2445
 Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met
 2450 2455 2460
 Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn
 2465 2470 2475 2480
 Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn
 2485 2490 2495
 Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro
 2500 2505 2510

-148-

Phe Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr
 2515 2520 2525
 Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg
 2530 2535 2540
 Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp
 2545 2550 2555 2560
 Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro
 2565 2570 2575
 Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val
 2580 2585 2590
 Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg
 2595 2600 2605
 Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile
 2610 2615 2620
 Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp
 2625 2630 2635 2640
 Thr Pro Tyr Met

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7624 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 333..7562

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTTGTGAAGA GAATGTTTTA CACTCTTGTT AGTGAAGTTT ATTCTTTAAA AGTCAATCGT	60
CAAGGATTTA GCAAATGAAT TAGCACTTCG GATATACTTG TTTATTTAAT ATCTTTTTTG	120
TTTATTTCAA AGAATTCAGT AATTGGATCA TAACGAGACT TCTGCGGATT GCAGCAACTC	180
CCTCCTGTCA TTTGTTACAC AAGAAAATCT GTGAAGTCAT CTGTTCATTA TTATTTCTTT	240
TTAAAAGCAA GAGTCCTGCT ATTTTGGGG TACTCACAAA AGAATTATTA CAACTTTTTG	300
AAGACTTGGT TTACCTCCAT AGAAGAAATG TG ATG GGT CAT GCT GTG GAA TGG	353
Met Gly His Ala Val Glu Trp	
1 5	
CCA GTG GTC ATG AGC CGA TTT TTA AGT CAA TTA GAT GAA CAC ATG GGA	401
Pro Val Val Met Ser Arg Phe Leu Ser Gln Leu Asp Glu His Met Gly	
10 15 20	
TAT TTA CAA TCA GCT CCT TTG CAG TTG ATG AGT ATG CAA AAT TTA GAA	449
Tyr Leu Gln Ser Ala Pro Leu Gln Leu Met Ser Met Gln Asn Leu Glu	
25 30 35	

-149-

TTT ATT GAA GTC ACT TTA TTA ATG GTT CTT ACT CGT ATT ATT GCA ATT Phe Ile Glu Val Thr Leu Leu Met Val Leu Thr Arg Ile Ile Ala Ile 40 45 50 55	497
GTG TTT TTT AGA AGG CAA GAA CTC TTA CTT TGG CAG ATA GGT TGT GTT Val Phe Phe Arg Arg Gln Glu Leu Leu Leu Trp Gln Ile Gly Cys Val 60 65 70	545
CTG CTA GAG TAT GGT AGT CCA AAA ATT AAA TCC CTA GCA ATT AGC TTT Leu Leu Glu Tyr Gly Ser Pro Lys Ile Lys Ser Leu Ala Ile Ser Phe 75 80 85	593
TTA ACA GAA CTT TTT CAG CTT GGA GGA CTA CCA GCA CAA CCA GCT AGC Leu Thr Glu Leu Phe Gln Leu Gly Gly Leu Pro Ala Gln Pro Ala Ser 90 95 100	641
ACT TTT TTC AGC TCA TTT TTG GAA TTA TTA AAA CAC CTT GTA GAA ATG Thr Phe Phe Ser Ser Phe Leu Glu Leu Leu Lys His Leu Val Glu Met 105 110 115	689
GAT ACT GAC CAA TTG AAA CTC TAT GAA GAG CCA TTA TCA AAG CTG ATA Asp Thr Asp Gln Leu Lys Leu Tyr Glu Glu Pro Leu Ser Lys Leu Ile 120 125 130 135	737
AAG ACA CTA TTT CCC TTT GAA GCA GAA GCT TAT AGA AAT ATT GAA CCT Lys Thr Leu Phe Pro Phe Glu Ala Glu Ala Tyr Arg Asn Ile Glu Pro 140 145 150	785
GTC TAT TTA AAT ATG CTG CTG GAA AAA CTC TGT GTC ATG TTT GAA GAC Val Tyr Leu Asn Met Leu Leu Glu Lys Leu Cys Val Met Phe Glu Asp 155 160 165	833
GGT GTG CTC ATG CGG CTT AAG TCT GAT TTG CTA AAA GCA GCT TTG TGC Gly Val Leu Met Arg Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys 170 175 180	881
CAT TTA CTG CAG TAT TTC CTT AAA TTT GTG CCA GCT GGG TAT GAA TCT His Leu Leu Gln Tyr Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser 185 190 195	929
GCT TTA CAA GTC AGG AAG GTC TAT GTG AGA AAT ATT TGT AAA GCT CTT Ala Leu Gln Val Arg Lys Val Tyr Val Arg Asn Ile Cys Lys Ala Leu 200 205 210 215	977
TTG GAT GTG CTT GGA ATT GAG GTA GAT GCA GAG TAC TTG TTG GGC CCA Leu Asp Val Leu Gly Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro 220 225 230	1025
CTT TAT GCA GCT TTG AAA ATG GAA AGT ATG GAA ATC ATT GAG GAG ATT Leu Tyr Ala Ala Leu Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile 235 240 245	1073
CAA TGC CAA ACT CAA CAG GAA AAC CTC AGC AGT AAT AGT GAT GGA ATA Gln Cys Gln Thr Gln Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile 250 255 260	1121
TCA CCC AAA AGG CGT CGT CTC AGC TCG TCT CTA AAC CCT TCT AAA AGA Ser Pro Lys Arg Arg Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg 265 270 275	1169
GCA CCA AAA CAG ACT GAG GAA ATT AAA CAT GTG GAC ATG AAC CAA AAG Ala Pro Lys Gln Thr Glu Glu Ile Lys His Val Asp Met Asn Gln Lys 280 285 290 295	1217
AGC ATA TTA TGG AGT GCA CTG AAA CAG AAA GCT GAA TCC CTT CAG ATT Ser Ile Leu Trp Ser Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile 300 305 310	1265

-150-

TCC	CTT	GAA	TAC	AGT	GGC	CTA	AAG	AAT	CCT	GTT	ATT	GAG	ATG	TTA	GAA	1313
Ser	Leu	Glu	Tyr	Ser	Gly	Leu	Lys	Asn	Pro	Val	Ile	Glu	Met	Leu	Glu	
		315						320					325			
GGA	ATT	GCT	GTT	GTC	TTA	CAA	CTG	ACT	GCT	CTG	TGT	ACT	GTT	CAT	TGT	1361
Gly	Ile	Ala	Val	Val	Leu	Gln	Leu	Thr	Ala	Leu	Cys	Thr	Val	His	Cys	
		330					335					340				
TCT	CAT	CAA	AAC	ATG	AAC	TGC	CGT	ACT	TTC	AAG	GAC	TGT	CAA	CAT	AAA	1409
Ser	His	Gln	Asn	Met	Asn	Cys	Arg	Thr	Phe	Lys	Asp	Cys	Gln	His	Lys	
		345				350					355					
TCC	AAG	AAG	AAA	CCT	TCT	GTA	GTG	ATA	ACT	TGG	ATG	TCA	TTG	GAT	TTT	1457
Ser	Lys	Lys	Lys	Pro	Ser	Val	Val	Ile	Thr	Trp	Met	Ser	Leu	Asp	Phe	
						365				370					375	
TAC	ACA	AAA	GTG	CTT	AAG	AGC	TGT	AGA	AGT	TTG	TTA	GAA	TCT	GTT	CAG	1505
Tyr	Thr	Lys	Val	Leu	Lys	Ser	Cys	Arg	Ser	Leu	Leu	Glu	Ser	Val	Gln	
				380					385					390		
AAA	CTG	GAC	CTG	GAG	GCA	ACC	ATT	GAT	AAG	GTG	GTG	AAA	ATT	TAT	GAT	1553
Lys	Leu	Asp	Leu	Glu	Ala	Thr	Ile	Asp	Lys	Val	Val	Lys	Ile	Tyr	Asp	
			395					400					405			
GCT	TTG	ATT	TAT	ATG	CAA	GTA	AAC	AGT	TCA	TTT	GAA	GAT	CAT	ATC	CTG	1601
Ala	Leu	Ile	Tyr	Met	Gln	Val	Asn	Ser	Ser	Phe	Glu	Asp	His	Ile	Leu	
		410					415					420				
GAA	GAT	TTA	TGT	GGA	ATG	CTC	TCA	CTT	CCA	TGG	ATT	TAT	TCC	CAT	TCT	1649
Glu	Asp	Leu	Cys	Gly	Met	Leu	Ser	Leu	Pro	Trp	Ile	Tyr	Ser	His	Ser	
		425				430					435					
GAT	GAT	GGC	TGT	TTA	AAG	TTG	ACC	ACA	TTT	GCC	GCT	AAT	CTT	CTA	ACA	1697
Asp	Asp	Gly	Cys	Leu	Lys	Leu	Thr	Thr	Phe	Ala	Ala	Asn	Leu	Leu	Thr	
					445					450					455	
TTA	AGC	TGT	AGG	ATT	TCA	GAT	AGC	TAT	TCA	CCA	CAG	GCA	CAA	TCA	CGA	1745
Leu	Ser	Cys	Arg	Ile	Ser	Asp	Ser	Tyr	Ser	Pro	Gln	Ala	Gln	Ser	Arg	
				460					465					470		
TGT	GTG	TTT	CTT	CTG	ACT	CTG	TTT	CCA	AGA	AGA	ATA	TTC	CTT	GAG	TGG	1793
Cys	Val	Phe	Leu	Leu	Thr	Leu	Phe	Pro	Arg	Arg	Ile	Phe	Leu	Glu	Trp	
			475					480					485			
AGA	ACA	GCA	GTT	TAC	AAC	TGG	GCC	CTG	CAG	AGC	TCC	CAT	GAA	GTA	ATC	1841
Arg	Thr	Ala	Val	Tyr	Asn	Trp	Ala	Leu	Gln	Ser	Ser	His	Glu	Val	Ile	
		490					495					500				
CGG	GCT	AGT	TGT	GTT	AGT	GGA	TTT	TTT	ATC	TTA	TTG	CAG	CAG	CAG	AAT	1889
Arg	Ala	Ser	Cys	Val	Ser	Gly	Phe	Phe	Ile	Leu	Leu	Gln	Gln	Gln	Asn	
		505				510					515					
TCT	TGT	AAC	AGA	GTT	CCC	AAG	ATT	CTT	ATA	GAT	AAA	GTC	AAA	GAT	GAT	1937
Ser	Cys	Asn	Arg	Val	Pro	Lys	Ile	Leu	Ile	Asp	Lys	Val	Lys	Asp	Asp	
					525					530					535	
TCT	GAC	ATT	GTC	AAG	AAA	GAA	TTT	GCT	TCT	ATA	CTT	GGT	CAA	CTT	GTC	1985
Ser	Asp	Ile	Val	Lys	Lys	Glu	Phe	Ala	Ser	Ile	Leu	Gly	Gln	Leu	Val	
				540					545					550		
TGT	ACT	CTT	CAC	GGC	ATG	TTT	TAT	CTG	ACA	AGT	TCT	TTA	ACA	GAA	CCT	2033
Cys	Thr	Leu	His	Gly	Met	Phe	Tyr	Leu	Thr	Ser	Ser	Leu	Thr	Glu	Pro	
				555				560					565			
TTC	TCT	GAA	CAC	GGA	CAT	GTG	GAC	CTC	TTC	TGT	AGG	AAC	TTG	AAA	GCC	2081
Phe	Ser	Glu	His	Gly	His	Val	Asp	Leu	Phe	Cys	Arg	Asn	Leu	Lys	Ala	
		570					575					580				

-151-

ACT Thr 585	TCT Ser 585	CAA Gln 585	CAT His 585	GAA Glu 585	TGT Cys 585	TCA Ser 590	TCT Ser 590	TCT Ser 590	CAA Gln 590	CTA Leu 595	AAA Lys 595	GCT Ala 595	TCT Ser 595	GTC Val 595	TGC Cys 595	2129
AAG Lys 600	CCA Pro 600	TTC Phe 600	CTT Leu 600	TTC Phe 605	CTA Leu 605	CTG Leu 605	AAA Lys 605	AAA Lys 605	AAA Lys 610	ATA Ile 610	CCT Pro 610	AGT Ser 610	CCA Pro 610	GTA Val 615	AAA Lys 615	2177
CTT Leu 620	GCT Ala 620	TTC Phe 620	ATA Ile 620	GAT Asp 620	AAT Asn 620	CTA Leu 625	CAT His 625	CAT His 625	CTT Leu 625	TGT Cys 625	AAG Lys 625	CAT His 625	CTT Leu 630	GAT Asp 630	TTT Phe 630	2225
AGA Arg 635	GAA Glu 635	GAT Asp 635	GAA Glu 635	ACA Thr 635	GAT Asp 635	GTA Val 640	AAA Lys 640	GCA Ala 640	GTT Val 640	CTT Leu 645	GGA Gly 645	ACT Thr 645	TTA Leu 645	TTA Leu 645	AAT Asn 645	2273
TTA Leu 650	ATG Met 650	GAA Glu 650	GAT Asp 650	CCA Pro 650	GAC Asp 655	AAA Lys 655	GAT Asp 655	GTT Val 655	AGA Arg 655	GTG Val 660	GCT Ala 660	TTT Phe 660	AGT Ser 660	GGA Gly 660	AAT Asn 660	2321
ATC Ile 665	AAG Lys 665	CAC His 665	ATA Ile 665	TTG Leu 665	GAA Glu 670	TCC Ser 670	TTG Leu 670	GAC Asp 670	TCT Ser 670	GAA Glu 675	GAT Asp 675	GGA Gly 675	TTT Phe 675	ATA Ile 675	AAG Lys 675	2369
GAG Glu 680	CTT Leu 680	TTT Phe 680	GTC Val 680	TTA Leu 685	AGA Arg 685	ATG Met 685	AAG Lys 685	GAA Glu 690	GCA Ala 690	TAT Tyr 690	ACA Thr 690	CAT His 690	GCC Ala 690	CAA Gln 695	ATA Ile 695	2417
TCA Ser 700	AGA Arg 700	AAT Asn 700	AAT Asn 700	GAG Glu 700	CTG Leu 700	AAG Lys 705	GAT Asp 705	ACC Thr 705	TTG Leu 705	ATT Ile 705	CTT Leu 710	ACA Thr 710	ACA Thr 710	GGG Gly 710	GAT Asp 710	2465
ATT Ile 715	GGA Gly 715	AGG Arg 715	GCC Ala 715	GCA Ala 715	AAA Lys 715	GGA Gly 720	GAT Asp 720	TTG Leu 720	GTA Val 720	CCA Pro 720	TTT Phe 725	GCA Ala 725	CTC Leu 725	TTA Leu 725	CAC His 725	2513
TTA Leu 730	TTG Leu 730	CAT His 730	TGT Cys 730	TTG Leu 735	TTA Leu 735	TCC Ser 735	AAG Lys 735	TCA Ser 735	GCA Ala 735	TCT Ser 740	GTC Val 740	TCT Ser 740	GGA Gly 740	GCA Ala 740	GCA Ala 740	2561
TAC Tyr 745	ACA Thr 745	GAA Glu 745	ATT Ile 745	AGA Arg 750	GCT Ala 750	CTG Leu 750	GTT Val 750	GCA Ala 750	GCT Ala 750	AAA Lys 755	AGT Ser 755	GTT Val 755	AAA Lys 755	CTG Leu 755	CAA Gln 755	2609
AGT Ser 760	TTT Phe 760	TTC Phe 760	AGC Ser 765	CAG Gln 765	TAT Tyr 765	AAG Lys 765	AAA Lys 765	CCC Pro 770	ATC Ile 770	TGT Cys 770	CAG Gln 770	TTT Phe 770	TTG Leu 775	GTA Val 775	GAA Glu 775	2657
TCC Ser 780	CTT Leu 780	CAC His 780	TCT Ser 780	AGT Ser 780	CAG Gln 785	ATG Met 785	ACA Thr 785	GCA Ala 785	CTT Leu 785	CCG Pro 785	AAT Asn 790	ACT Thr 790	CCA Pro 790	TGC Cys 790	CAG Gln 790	2705
AAT Asn 795	GCT Ala 795	GAC Asp 795	GTG Val 795	CGA Arg 795	AAA Lys 800	CAA Gln 800	GAT Asp 800	GTG Val 800	GCT Ala 805	CAC His 805	CAG Gln 805	AGA Arg 805	GAA Glu 805	ATG Met 805	GCT Ala 805	2753
TTA Leu 810	AAT Asn 810	ACG Thr 810	TTG Leu 810	TCT Ser 815	GAA Glu 815	ATT Ile 815	GCC Ala 815	AAC Asn 815	GTT Val 815	TTC Phe 820	GAC Asp 820	TTT Phe 820	CCT Pro 820	GAT Asp 820	CTT Leu 820	2801
AAT Asn 825	CGT Arg 825	TTT Phe 825	CTT Leu 825	ACT Thr 830	AGG Arg 830	ACA Thr 830	TTA Leu 830	CAA Gln 830	GTT Val 830	CTA Leu 835	CTA Leu 835	CCT Pro 835	GAT Asp 835	CTT Leu 835	GCT Ala 835	2849
GCC Ala 840	AAA Lys 840	GCA Ala 840	AGC Ser 840	CCT Pro 845	GCA Ala 845	GCT Ala 845	TCT Ser 845	GCT Ala 845	CTC Leu 850	ATT Ile 850	CGA Arg 850	ACT Thr 850	TTA Leu 850	GGA Gly 850	AAA Lys 855	2897

-152-

CAA TTA AAT GTC AAT CGT AGA GAG ATT TTA ATA AAC AAC TTC AAA TAT	2945
Gln Leu Asn Val Asn Arg Arg Glu Ile Leu Ile Asn Asn Phe Lys Tyr	
860 865 870	
ATT TTT TCT CAT TTG GTC TGT TCT TGT TCC AAA GAT GAA TTA GAA CGT	2993
Ile Phe Ser His Leu Val Cys Ser Cys Ser Lys Asp Glu Leu Glu Arg	
875 880 885	
GCC CTT CAT TAT CTG AAG AAT GAA ACA GAA ATT GAA CTG GGG AGC CTG	3041
Ala Leu His Tyr Leu Lys Asn Glu Thr Glu Ile Glu Leu Gly Ser Leu	
890 895 900	
TTG AGA CAA GAT TTC CAA GGA TTG CAT AAT GAA TTA TTG CTG CGT ATT	3089
Leu Arg Gln Asp Phe Gln Gly Leu His Asn Glu Leu Leu Leu Arg Ile	
905 910 915	
GGA GAA CAC TAT CAA CAG GTT TTT AAT GGT TTG TCA ATA CTT GCC TCA	3137
Gly Glu His Tyr Gln Gln Val Phe Asn Gly Leu Ser Ile Leu Ala Ser	
920 925 930 935	
TTT GCA TCC AGT GAT GAT CCA TAT CAG GGC CCG AGA GAT ATC ATA TCA	3185
Phe Ala Ser Ser Asp Asp Pro Tyr Gln Gly Pro Arg Asp Ile Ile Ser	
940 945 950	
CCT GAA CTG ATG GCT GAT TAT TTA CAA CCC AAA TTG TTG GGC ATT TTG	3233
Pro Glu Leu Met Ala Asp Tyr Leu Gln Pro Lys Leu Leu Gly Ile Leu	
955 960 965	
GCT TTT TTT AAC ATG CAG TTA CTG AGC TCT AGT GTT GGC ATT GAA GAT	3281
Ala Phe Phe Asn Met Gln Leu Leu Ser Ser Ser Val Gly Ile Glu Asp	
970 975 980	
AAG AAA ATG GCC TTG AAC AGT TTG ATG TCT TTG ATG AAG TTA ATG GGA	3329
Lys Lys Met Ala Leu Asn Ser Leu Met Ser Leu Met Lys Leu Met Gly	
985 990 995	
CCC AAA CAT GTC AGT TCT GTG AGG GTG AAG ATG ATG ACC ACA CTG AGA	3377
Pro Lys His Val Ser Ser Val Arg Val Lys Met Met Thr Thr Leu Arg	
1000 1005 1010 1015	
ACT GGC CTT CGA TTC AAG GAT GAT TTT CCT GAA TTG TGT TGC AGA GCT	3425
Thr Gly Leu Arg Phe Lys Asp Asp Phe Pro Glu Leu Cys Cys Arg Ala	
1020 1025 1030	
TGG GAC TGC TTT GTT CGC TGC CTG GAT CAT GCT TGT CTG GGC TCC CTT	3473
Trp Asp Cys Phe Val Arg Cys Leu Asp His Ala Cys Leu Gly Ser Leu	
1035 1040 1045	
CTC AGT CAT GTA ATA GTA GCT TTG TTA CCT CTT ATA CAC ATC CAG CCT	3521
Leu Ser His Val Ile Val Ala Leu Leu Pro Leu Ile His Ile Gln Pro	
1050 1055 1060	
AAA GAA ACT GCA GCT ATC TTC CAC TAC CTC ATA ATT GAA AAC AGG GAT	3569
Lys Glu Thr Ala Ala Ile Phe His Tyr Leu Ile Ile Glu Asn Arg Asp	
1065 1070 1075	
GCT GTG CAA GAT TTT CTT CAT GAA ATA TAT TTT TTA CCT GAT CAT CCA	3617
Ala Val Gln Asp Phe Leu His Glu Ile Tyr Phe Leu Pro Asp His Pro	
1080 1085 1090 1095	
GAA TTA AAA AAG ATA AAA GCC GTT CTC CAG GAA TAC AGA AAG GAG ACC	3665
Glu Leu Lys Lys Ile Lys Ala Val Leu Gln Glu Tyr Arg Lys Glu Thr	
1100 1105 1110	
TCT GAG AGC ACT GAT CTT CAG ACA ACT CTT CAG CTC TCT ATG AAG GCC	3713
Ser Glu Ser Thr Asp Leu Gln Thr Thr Leu Gln Leu Ser Met Lys Ala	
1115 1120 1125	

-153-

ATT CAA CAT GAA AAT GTC GAT GTT CGT ATT CAT GCT CTT ACA AGC TTG Ile Gln His Glu Asn Val Asp Val Arg Ile His Ala Leu Thr Ser Leu 1130 1135 1140	3761
AAG GAA ACC TTG TAT AAA AAT CAG GAA AAA CTG ATA AAG TAT GCA ACA Lys Glu Thr Leu Tyr Lys Asn Gln Glu Lys Leu Ile Lys Tyr Ala Thr 1145 1150 1155	3809
GAC AGT GAA ACA GTA GAA CCT ATT ATC TCA CAG TTG GTG ACA GTG CTT Asp Ser Glu Thr Val Glu Pro Ile Ile Ser Gln Leu Val Thr Val Leu 1160 1165 1170 1175	3857
TTG AAA GGT TGC CAA GAT GCA AAC TCT CAA GCT CGG TTG CTC TGT GGG Leu Lys Gly Cys Gln Asp Ala Asn Ser Gln Ala Arg Leu Leu Cys Gly 1180 1185 1190	3905
GAA TGT TTA GGG GAA TTG GGG GCG ATA GAT CCA GGT CGA TTA GAT TTC Glu Cys Leu Gly Glu Leu Gly Ala Ile Asp Pro Gly Arg Leu Asp Phe 1195 1200 1205	3953
TCA ACA ACT GAA ACT CAA GGA AAA GAT TTT ACA TTT GTG ACT GGA GTA Ser Thr Thr Glu Thr Gln Gly Lys Asp Phe Thr Phe Val Thr Gly Val 1210 1215 1220	4001
GAA GAT TCA AGC TTT GCC TAT GGA TTA TTG ATG GAG CTA ACA AGA GCT Glu Asp Ser Ser Phe Ala Tyr Gly Leu Leu Met Glu Leu Thr Arg Ala 1225 1230 1235	4049
TAC CTT GCG TAT GCT GAT AAT AGC CGA GCT CAA GAT TCA GCT GCC TAT Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala Gln Asp Ser Ala Ala Tyr 1240 1245 1250 1255	4097
GCC ATT CAG GAG TTG CTT TCT ATT TAT GAC TGT AGA GAG ATG GAG ACC Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp Cys Arg Glu Met Glu Thr 1260 1265 1270	4145
AAC GGC CCA GGT CAC CAA TTG TGG AGG AGA TTT CCT GAG CAT GTT CGG Asn Gly Pro Gly His Gln Leu Trp Arg Arg Phe Pro Glu His Val Arg 1275 1280 1285	4193
GAA ATA CTA GAA CCT CAT CTA AAT ACC AGA TAC AAG AGT TCT CAG AAG Glu Ile Leu Glu Pro His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys 1290 1295 1300	4241
TCA ACC GAT TGG TCT GGA GTA AAG AAG CCA ATT TAC TTA AGT AAA TTG Ser Thr Asp Trp Ser Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu 1305 1310 1315	4289
GGT AGT AAC TTT GCA GAA TGG TCA GCA TCT TGG GCA GGT TAT CTT ATT Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile 1320 1325 1330 1335	4337
ACA AAG GTT CGA CAT GAT CTT GCC AGT AAA ATT TTC ACC TGC TGT AGC Thr Lys Val Arg His Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser 1340 1345 1350	4385
ATT ATG ATG AAG CAT GAT TTC AAA GTG ACC ATC TAT CTT CTT CCA CAT Ile Met Met Lys His Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His 1355 1360 1365	4433
ATT CTG GTG TAT GTC TTA CTG GGT TGT AAT CAA GAA GAT CAG CAG GAG Ile Leu Val Tyr Val Leu Leu Gly Cys Asn Gln Glu Asp Gln Gln Glu 1370 1375 1380	4481
GTT TAT GCA GAA ATT ATG GCA GTT CTA AAG CAT GAC GAT CAG CAT ACC Val Tyr Ala Glu Ile Met Ala Val Leu Lys His Asp Asp Gln His Thr 1385 1390 1395	4529

-154-

ATA AAT ACC CAA GAC ATT GCA TCT GAT CTG TGT CAA CTC AGT ACA CAG Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln 1400 1405 1410 1415	4577
ACT GTG TTC TCC ATG CTT GAC CAT CTC ACA CAG TGG GCA AGG CAC AAA Thr Val Phe Ser Met Leu Asp His Leu Thr Gln Trp Ala Arg His Lys 1420 1425 1430	4625
TTT CAG GCA CTG AAA GCT GAG AAA TGT CCA CAC AGC AAA TCA AAC AGA Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg 1435 1440 1445	4673
AAT AAG GTA GAC TCA ATG GTA TCT ACT GTG GAT TAT GAA GAC TAT CAG Asn Lys Val Asp Ser Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln 1450 1455 1460	4721
AGT GTA ACC CGT TTT CTA GAC CTC ATA CCC CAG GAT ACT CTG GCA GTA Ser Val Thr Arg Phe Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val 1465 1470 1475	4769
GCT TCC TTT CGC TCC AAA GCA TAC ACA CGA GCT GTA ATG CAC TTT GAA Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu 1480 1485 1490 1495	4817
TCA TTT ATT ACA GAA AAG AAG CAA AAT ATT CAG GAA CAT CTT GGA TTT Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile Gln Glu His Leu Gly Phe 1500 1505 1510	4865
TTA CAG AAA TTG TAT GCT GCT ATG CAT GAA CCT GAT GGA GTG GCC GGA Leu Gln Lys Leu Tyr Ala Ala Met His Glu Pro Asp Gly Val Ala Gly 1515 1520 1525	4913
GTC AGT GCA ATT AGA AAG GCA GAA CCA TCT CTA AAA GAA CAG ATC CTT Val Ser Ala Ile Arg Lys Ala Glu Pro Ser Leu Lys Glu Gln Ile Leu 1530 1535 1540	4961
GAA CAT GAA AGC CTT GGC TTG CTG AGG GAT GCC ACT GCT TGT TAT GAC Glu His Glu Ser Leu Gly Leu Leu Arg Asp Ala Thr Ala Cys Tyr Asp 1545 1550 1555	5009
AGG GCT ATT CAG CTA GAA CCA GAC CAG ATC ATT CAT TAC CAT GGT GTA Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile Ile His Tyr His Gly Val 1560 1565 1570 1575	5057
GTA AAG TCC ATG TTA GGT CTT GGT CAG CTG TCT ACT GTT ATC ACT CAG Val Lys Ser Met Leu Gly Leu Gly Gln Leu Ser Thr Val Ile Thr Gln 1580 1585 1590	5105
GTG AAT GGA GTG CAT GCT AAC AGG TCC GAG TGG ACA GAT GAA TTA AAC Val Asn Gly Val His Ala Asn Arg Ser Glu Trp Thr Asp Glu Leu Asn 1595 1600 1605	5153
ACG TAC AGA GTG GAA GCA GCT TGG AAA TTG TCA CAG TGG GAT TTG GTG Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val 1610 1615 1620	5201
GAA AAC TAT TTG GCA GCA GAT GGA AAA TCT ACA ACA TGG AGT GTC AGA Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg 1625 1630 1635	5249
CTG GGA CAG CTA TTA TTA TCA GCC AAA AAA AGA GAT ATC ACA GCT TTT Leu Gly Gln Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe 1640 1645 1650 1655	5297
TAT GAC TCA CTG AAA CTA GTG AGA GCA GAA CAA ATT GTA CCT CTT TCA Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser 1660 1665 1670	5345

-155-

GCT GCA AGC TTT GAA AGA GGC TCC TAC CAA CGA GGA TAT GAA TAT ATT	5393
Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile	
1675 1680 1685	
GTG AGA TTG CAC ATG TTA TGT GAG TTG GAG CAT AGC ATC AAA CCA CTT	5441
Val Arg Leu His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu	
1690 1695 1700	
TTC CAG CAT TCT CCA GGT GAC AGT TCT CAA GAA GAT TCT CTA AAC TGG	5489
Phe Gln His Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp	
1705 1710 1715	
GTA GCT CGA CTA GAA ATG ACC CAG AAT TCC TAC AGA GCC AAG GAG CCT	5537
Val Ala Arg Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro	
1720 1725 1730 1735	
ATC CTG GCT CTC CGG AGG GCT TTA CTA AGC CTC AAC AAA AGA CCA GAT	5585
Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp	
1740 1745 1750	
TAC AAT GAA ATG GTT GGA GAA TGC TGG CTG CAG AGT GCC AGG GTA GCT	5633
Tyr Asn Glu Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala	
1755 1760 1765	
AGA AAG GCT GGT CAC CAC CAG ACA GCC TAC AAT GCT CTC CTT AAT GCA	5681
Arg Lys Ala Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala	
1770 1775 1780	
GGG GAA TCA CGA CTC GCT GAA CTG TAC GTG GAA AGG GCA AAG TGG CTC	5729
Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu	
1785 1790 1795	
TGG TCC AAG GGT GAT GTT CAC CAG GCA CTA ATT GTT CTT CAA AAA GGT	5777
Trp Ser Lys Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly	
1800 1805 1810 1815	
GTT GAA TTA TGT TTT CCT GAA AAT GAA ACC CCA CCT GAG GGT AAG AAC	5825
Val Glu Leu Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn	
1820 1825 1830	
ATG TTA ATC CAT GGT CGA GCT ATG CTA CTA GTG GGC CGA TTT ATG GAA	5873
Met Leu Ile His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu	
1835 1840 1845	
GAA ACA GCT AAC TTT GAA AGC AAT GCA ATT ATG AAA AAA TAT AAG GAT	5921
Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp	
1850 1855 1860	
GTG ACC GCG TGC CTG CCA GAA TGG GAG GAT GGG CAT TTT TAC CTT GCC	5969
Val Thr Ala Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala	
1865 1870 1875	
AAG TAC TAT GAC AAA TTG ATG CCC ATG GTC ACA GAC AAC AAA ATG GAA	6017
Lys Tyr Tyr Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu	
1880 1885 1890 1895	
AAG CAA GGT GAT CTC ATC CGG TAT ATA GTT CTT CAT TTT GGC AGA TCT	6065
Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser	
1900 1905 1910	
CTA CAA TAT GGA AAT CAG TTC ATA TAT CAG TCA ATG CCA CGA ATG TTA	6113
Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu	
1915 1920 1925	
ACT CTA TGG CTT GAT TAT GGT ACA AAG GCA TAT GAA TGG GAA AAA GCT	6161
Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala	
1930 1935 1940	

-156-

GGC CGC TCC GAT CGT GTA CAA ATG AGG AAT GAT TTG GGT AAA ATA AAC Gly Arg Ser Asp Arg Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn 1945 1950 1955	6209
AAG GTT ATC ACA GAG CAT ACA AAC TAT TTA GCT CCA TAT CAA TTT TTG Lys Val Ile Thr Glu His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu 1960 1965 1970 1975	6257
ACT GCT TTT TCA CAA TTG ATC TCT CGA ATT TGT CAT TCT CAC GAT GAA Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile Cys His Ser His Asp Glu 1980 1985 1990	6305
GTT TTT GTT GTC TTG ATG GAA ATA ATA GCC AAA GTA TTT CTA GCC TAT Val Phe Val Val Leu Met Glu Ile Ile Ala Lys Val Phe Leu Ala Tyr 1995 2000 2005	6353
CCT CAA CAA GCA ATG TGG ATG ATG ACA GCT GTG TCA AAG TCA TCT TAT Pro Gln Gln Ala Met Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr 2010 2015 2020	6401
CCC ATG CGT GTG AAC AGA TGC AAG GAA ATC CTC AAT AAA GCT ATT CAT Pro Met Arg Val Asn Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His 2025 2030 2035	6449
ATG AAA AAA TCC TTA GAG AAG TTT GTT GGA GAT GCA ACT CGC CTA ACA Met Lys Lys Ser Leu Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr 2040 2045 2050 2055	6497
GAT AAG CTT CTA GAA TTG TGC AAT AAA CCG GTG GAA ATT CTT GCT TCT Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro Val Glu Ile Leu Ala Ser 2060 2065 2070	6545
CTT CAG AAA CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe 2075 2080 2085	6593
TAC ATC ATG ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg 2090 2095 2100	6641
CTA ATG GAA TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala 2105 2110 2115	6689
GAG TCT CGT AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA Glu Ser Arg Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro 2120 2125 2130 2135	6737
CTA AAT GAT GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly 2140 2145 2150	6785
TTG AGA CCT ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met 2155 2160 2165	6833
ACA GGA AAA GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu 2170 2175 2180	6881
TCT GAA AAA CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro 2185 2190 2195	6929
CCT ATT TTT CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser 2200 2205 2210 2215	6977

-157-

TGG TAC AGT AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser 2220 2225 2230	7025
ATG GTT GGT TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile 2235 2240 2245	7073
CTC TTT GAT TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys 2250 2255 2260	7121
CTT TTC AAT AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe 2265 2270 2275	7169
CGC CTG ACT CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu 2280 2285 2290 2295	7217
GGT CTT TTT CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp 2300 2305 2310	7265
CAG CGA GAG CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro 2315 2320 2325	7313
CTT GTG GAA TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu 2330 2335 2340	7361
AAT GAA ACT GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu 2345 2350 2355	7409
GAC ATT GAG CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val 2360 2365 2370 2375	7457
ACA GGA CTG CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln 2380 2385 2390	7505
GAA GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr 2395 2400 2405	7553
CCA TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA Pro Tyr Met 2410	7602
AGTAAAAAAA AAAAAAAAAA AA	7624

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2410 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

-158-

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser
 1 5 10 15
 Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu
 20 25 30
 Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val
 35 40 45
 Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu
 50 55 60
 Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile
 65 70 75 80
 Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly
 85 90 95
 Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu
 100 105 110
 Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu
 115 120 125
 Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu
 130 135 140
 Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys
 145 150 155 160
 Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp
 165 170 175
 Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe
 180 185 190
 Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val
 195 200 205
 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp
 210 215 220
 Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser
 225 230 235 240
 Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu
 245 250 255
 Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350

-159-

Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg
 370 375 380
 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp
 385 390 395 400
 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser
 405 410 415
 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu
 420 425 430
 Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr
 435 440 445
 Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr
 450 455 460
 Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro
 465 470 475 480
 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu
 485 490 495
 Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe
 500 505 510
 Ile Leu Leu Gln Gln Gln Asn Ser Cys Asn Arg Val Pro Lys Ile Leu
 515 520 525
 Ile Asp Lys Val Lys Asp Asp Ser Asp Ile Val Lys Lys Glu Phe Ala
 530 535 540
 Ser Ile Leu Gly Gln Leu Val Cys Thr Leu His Gly Met Phe Tyr Leu
 545 550 555 560
 Thr Ser Ser Leu Thr Glu Pro Phe Ser Glu His Gly His Val Asp Leu
 565 570 575
 Phe Cys Arg Asn Leu Lys Ala Thr Ser Gln His Glu Cys Ser Ser Ser
 580 585 590
 Gln Leu Lys Ala Ser Val Cys Lys Pro Phe Leu Phe Leu Leu Lys Lys
 595 600 605
 Lys Ile Pro Ser Pro Val Lys Leu Ala Phe Ile Asp Asn Leu His His
 610 615 620
 Leu Cys Lys His Leu Asp Phe Arg Glu Asp Glu Thr Asp Val Lys Ala
 625 630 635 640
 Val Leu Gly Thr Leu Leu Asn Leu Met Glu Asp Pro Asp Lys Asp Val
 645 650 655
 Arg Val Ala Phe Ser Gly Asn Ile Lys His Ile Leu Glu Ser Leu Asp
 660 665 670
 Ser Glu Asp Gly Phe Ile Lys Glu Leu Phe Val Leu Arg Met Lys Glu
 675 680 685
 Ala Tyr Thr His Ala Gln Ile Ser Arg Asn Asn Glu Leu Lys Asp Thr
 690 695 700

-160-

Leu Ile Leu Thr Thr Gly Asp Ile Gly Arg Ala Ala Lys Gly Asp Leu
 705 710 715 720
 Val Pro Phe Ala Leu Leu His Leu Leu His Cys Leu Leu Ser Lys Ser
 725 730 735
 Ala Ser Val Ser Gly Ala Ala Tyr Thr Glu Ile Arg Ala Leu Val Ala
 740 745 750
 Ala Lys Ser Val Lys Leu Gln Ser Phe Phe Ser Gln Tyr Lys Lys Pro
 755 760 765
 Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala
 770 775 780
 Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val
 785 790 795 800
 Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn
 805 810 815
 Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln
 820 825 830
 Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala
 835 840 845
 Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile
 850 855 860
 Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys
 865 870 875 880
 Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr
 885 890 895
 Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His
 900 905 910
 Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn
 915 920 925
 Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln
 930 935 940
 Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln
 945 950 955 960
 Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser
 965 970 975
 Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met
 980 985 990
 Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val
 995 1000 1005
 Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe
 1010 1015 1020
 Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp
 1025 1030 1035 1040
 His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu
 1045 1050 1055

-161-

Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr
 1060 1065 1070
 Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile
 1075 1080 1085
 Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu
 1090 1095 1100
 Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr
 1105 1110 1115 1120
 Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg
 1125 1130 1135
 Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu
 1140 1145 1150
 Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile
 1155 1160 1165
 Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser
 1170 1175 1180
 Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile
 1185 1190 1195 1200
 Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp
 1205 1210 1215
 Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu
 1220 1225 1230
 Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg
 1235 1240 1245
 Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr
 1250 1255 1260
 Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg
 1265 1270 1275 1280
 Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr
 1285 1290 1295
 Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys
 1300 1305 1310
 Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala
 1315 1320 1325
 Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser
 1330 1335 1340
 Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val
 1345 1350 1355 1360
 Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys
 1365 1370 1375
 Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu
 1380 1385 1390
 Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp
 1395 1400 1405

-162-

Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu
 1410 1415 1420
 Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys
 1425 1430 1435 1440
 Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr
 1445 1450 1455
 Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile
 1460 1465 1470
 Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr
 1475 1480 1485
 Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn
 1490 1495 1500
 Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His
 1505 1510 1515 1520
 Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro
 1525 1530 1535
 Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg
 1540 1545 1550
 Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln
 1555 1560 1565
 Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln
 1570 1575 1580
 Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser
 1585 1590 1595 1600
 Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys
 1605 1610 1615
 Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys
 1620 1625 1630
 Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys
 1635 1640 1645
 Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala
 1650 1655 1660
 Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr
 1665 1670 1675 1680
 Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu
 1685 1690 1695
 Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser
 1700 1705 1710
 Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn
 1715 1720 1725
 Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu
 1730 1735 1740
 Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp
 1745 1750 1755 1760

-163-

Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala
 1765 1770 1775
 Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr
 1780 1785 1790
 Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala
 1795 1800 1805
 Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu
 1810 1815 1820
 Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu
 1825 1830 1835 1840
 Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala
 1845 1850 1855
 Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu
 1860 1865 1870
 Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met
 1875 1880 1885
 Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile
 1890 1895 1900
 Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr
 1905 1910 1915 1920
 Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys
 1925 1930 1935
 Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg
 1940 1945 1950
 Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr
 1955 1960 1965
 Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg
 1970 1975 1980
 Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Met Glu Ile Ile
 1985 1990 1995 2000
 Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr
 2005 2010 2015
 Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu
 2020 2025 2030
 Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val
 2035 2040 2045
 Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys
 2050 2055 2060
 Pro Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu
 2065 2070 2075 2080
 Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp
 2085 2090 2095
 Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn
 2100 2105 2110

-164-

Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile
 2115 2120 2125
 Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu
 2130 2135 2140
 Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr
 2145 2150 2155 2160
 Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met
 2165 2170 2175
 Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu
 2180 2185 2190
 Phe Leu Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg
 2195 2200 2205
 Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys
 2210 2215 2220
 Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly
 2225 2230 2235 2240
 Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys
 2245 2250 2255
 Val His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu
 2260 2265 2270
 Val Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly
 2275 2280 2285
 Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val
 2290 2295 2300
 Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu
 2305 2310 2315 2320
 Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys
 2325 2330 2335
 Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu
 2340 2345 2350
 Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val
 2355 2360 2365
 Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly
 2370 2375 2380
 His Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys
 2385 2390 2395 2400
 Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2405 2410

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7502 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

-165-

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..7440

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG GGT CAT GCT GTG GAA TGG CCA GTG GTC ATG AGC CGA TTT TTA AGT	48
Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser	
1 5 10 15	
CAA TTA GAT GAA CAC ATG GGA TAT TTA CAA TCA GCT CCT TTG CAG TTG	96
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu	
20 25 30	
ATG AGT ATG CAA AAT TTA GAA TTT ATT GAA GTC ACT TTA TTA ATG GTT	144
Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val	
35 40 45	
CTT ACT CGT ATT ATT GCA ATT GTG TTT TTT AGA AGG CAA GAA CTC TTA	192
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu	
50 55 60	
CTT TGG CAG ATA GGT TGT GTT CTG CTA GAG TAT GGT AGT CCA AAA ATT	240
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile	
65 70 75 80	
AAA TCC CTA GCA ATT AGC TTT TTA ACA GAA CTT TTT CAG CTT GGA GGA	288
Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly	
85 90 95	
CTA CCA GCA CAA CCA GCT AGC ACT TTT TTC AGC TCA TTT TTG GAA TTA	336
Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu	
100 105 110	
TTA AAA CAC CTT GTA GAA ATG GAT ACT GAC CAA TTG AAA CTC TAT GAA	384
Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu	
115 120 125	
GAG CCA TTA TCA AAG CTG ATA AAG ACA CTA TTT CCC TTT GAA GCA GAA	432
Glu Pro Leu Ser Lys Leu Lys Thr Leu Phe Pro Phe Glu Ala Glu	
130 135 140	
GCT TAT AGA AAT ATT GAA CCT GTC TAT TTA AAT ATG CTG CTG GAA AAA	480
Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys	
145 150 155 160	
CTC TGT GTC ATG TTT GAA GAC GGT GTG CTC ATG CGG CTT AAG TCT GAT	528
Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp	
165 170 175	
TTG CTA AAA GCA GCT TTG TGC CAT TTA CTG CAG TAT TTC CTT AAA TTT	576
Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe	
180 185 190	
GTG CCA GCT GGG TAT GAA TCT GCT TTA CAA GTC AGG AAG GTC TAT GTG	624
Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val	
195 200 205	
AGA AAT ATT TGT AAA GCT CTT TTG GAT GTG CTT GGA ATT GAG GTA GAT	672
Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp	
210 215 220	

-166-

GCA GAG TAC TTG TTG GGC CCA CTT TAT GCA GCT TTG AAA ATG GAA AGT Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser 225 230 235 240	720
ATG GAA ATC ATT GAG GAG ATT CAA TGC CAA ACT CAA CAG GAA AAC CTC Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu 245 250 255	768
AGC AGT AAT AGT GAT GGA ATA TCA CCC AAA AGG CGT CGT CTC AGC TCG Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser 260 265 270	816
TCT CTA AAC CCT TCT AAA AGA GCA CCA AAA CAG ACT GAG GAA ATT AAA Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys 275 280 285	864
CAT GTG GAC ATG AAC CAA AAG AGC ATA TTA TGG AGT GCA CTG AAA CAG His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln 290 295 300	912
AAA GCT GAA TCC CTT CAG ATT TCC CTT GAA TAC AGT GGC CTA AAG AAT Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn 305 310 315 320	960
CCT GTT ATT GAG ATG TTA GAA GGA ATT GCT GTT GTC TTA CAA CTG ACT Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr 325 330 335	1008
GCT CTG TGT ACT GTT CAT TGT TCT CAT CAA AAC ATG AAC TGC CGT ACT Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 340 345 350	1056
TTC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAA CCT TCT GTA GTG ATA Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile 355 360 365	1104
ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 370 375 380	1152
AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG GCA ACC ATT GAT Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 385 390 395 400	1200
AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser 405 410 415	1248
TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TGT GGA ATG CTC TCA CTT Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 425 430	1296
CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 440 445	1344
TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 455 460	1392
TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 465 470 475 480	1440
AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu 485 490 495	1488

-167-

CAG Gln	AGC Ser	TCC Ser	CAT His 500	GAA Glu	GTA Val	ATC Ile	CGG Arg 505	GCT Ala 505	AGT Ser	TGT Cys	GTT Val	AGT Ser	GGA Gly 510	TTT Phe	TTT Phe	1536
ATC Ile	TTA Leu	TTG Leu	CAG Gln 515	CAG Gln	CAG Gln	AAT Asn 520	TCT Ser	TGT Cys	AAC Asn	AGA Arg	GTT Val	CCC Pro 525	AAG Lys	ATT Ile	CTT Leu	1584
ATA Ile	GAT Asp 530	AAA Lys	GTC Val	AAA Lys	GAT Asp	GAT Asp 535	TCT Ser	GAC Asp	ATT Ile	GTC Val	AAG Lys 540	AAA Lys	GAA Glu	TTT Phe	GCT Ala	1632
TCT Ser 545	ATA Ile	CTT Leu	GGT Gly	CAA Gln	CTT Leu 550	GTC Val	TGT Cys	ACT Thr	CTT Leu	CAC His 555	GGC Gly	ATG Met	TTT Phe	TAT Tyr	CTG Leu 560	1680
ACA Thr	AGT Ser	TCT Ser	TTA Leu 565	ACA Thr	GAA Glu	CCT Pro	TTC Phe	TCT Ser	GAA Glu 570	CAC His	GGA Gly	CAT His	GTG Val	GAC Asp 575	CTC Leu	1728
TTC Phe	TGT Cys	AGG Arg	AAC Asn 580	TTG Leu	AAA Lys	GCC Ala	ACT Thr	TCT Ser 585	CAA Gln	CAT His	GAA Glu	TGT Cys	TCA Ser 590	TCT Ser	TCT Ser	1776
CAA Gln	CTA Leu	AAA Lys 595	GCT Ala	TCT Ser	GTC Val	TGC Cys	AAG Lys 600	CCA Pro	TTC Phe	CTT Leu	TTC Phe	CTA Leu 605	CTG Leu	AAA Lys	AAA Lys	1824
AAA Lys 610	ATA Ile	CCT Pro	AGT Ser	CCA Pro	GTA Val	AAA Lys 615	CTT Leu	GCT Ala	TTC Phe	ATA Ile	GAT Asp 620	AAT Asn	CTA Leu	CAT His	CAT His	1872
CTT Leu 625	TGT Cys	AAG Lys	CAT His	CTT Leu	GAT Asp 630	TTT Phe	AGA Arg	GAA Glu	GAT Asp	GAA Glu 635	ACA Thr	GAT Asp	GTA Val	AAA Lys	GCA Ala 640	1920
GTT Val	CTT Leu	GGA Gly	ACT Thr 645	TTA Leu	TTA Leu	AAT Asn	TTA Leu	ATG Met	GAA Glu 650	GAT Asp	CCA Pro	GAC Asp	AAA Lys	GAT Asp 655	GTT Val	1968
AGA Arg	GTG Val	GCT Ala	TTT Phe 660	AGT Ser	GGA Gly	AAT Asn	ATC Ile	AAG Lys 665	CAC His	ATA Ile	TTG Leu	GAA Glu	TCC Ser 670	TTG Leu	GAC Asp	2016
TCT Ser	GAA Glu	GAT Asp 675	GGA Gly	TTT Phe	ATA Ile	AAG Lys	GAG Glu 680	CTT Leu	TTT Phe	GTC Val	TTA Leu	AGA Arg 685	ATG Met	AAG Lys	GAA Glu	2064
GCA Ala	TAT Tyr 690	ACA Thr	CAT His	GCC Ala	CAA Gln	ATA Ile 695	TCA Ser	AGA Arg	AAT Asn	AAT Asn	GAG Glu 700	CTG Leu	AAG Lys	GAT Asp	ACC Thr	2112
TTG Leu 705	ATT Ile	CTT Leu	ACA Thr	ACA Thr	GGG Gly 710	GAT Asp	ATT Ile	GGA Gly	AGG Arg	GCC Ala 715	GCA Ala	AAA Lys	GGA Gly	GAT Asp	TTG Leu 720	2160
GTA Val	CCA Pro	TTT Phe	GCA Ala 725	CTC Leu	TTA Leu	CAC His	TTA Leu	TTG Leu 730	CAT His	TGT Cys	TTG Leu	TTA Leu	TCC Ser 735	AAG Lys	TCA Ser	2208
GCA Ala	TCT Ser	GTC Val	TCT Ser 740	GGA Gly	GCA Ala	GCA Ala	TAC Tyr 745	ACA Thr	GAA Glu	ATT Ile	AGA Arg	GCT Ala 750	CTG Leu	GTT Val	GCA Ala	2256
GCT Ala	AAA Lys 755	AGT Ser	GTT Val	AAA Lys	CTG Leu	CAA Gln	AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	2304

-168-

ATC TGT CAG TTT TTG GTA GAA TCC CTT CAC TCT AGT CAG ATG ACA GCA Ile Cys Gln Phe Leu Val Glu Ser Leu His Ser Ser Gln Met Thr Ala 770 775 780	2352
CTT CCG AAT ACT CCA TGC CAG AAT GCT GAC GTG CGA AAA CAA GAT GTG Leu Pro Asn Thr Pro Cys Gln Asn Ala Asp Val Arg Lys Gln Asp Val 785 790 795 800	2400
GCT CAC CAG AGA GAA ATG GCT TTA AAT ACG TTG TCT GAA ATT GCC AAC Ala His Gln Arg Glu Met Ala Leu Asn Thr Leu Ser Glu Ile Ala Asn 805 810 815	2448
GTT TTC GAC TTT CCT GAT CTT AAT CGT TTT CTT ACT AGG ACA TTA CAA Val Phe Asp Phe Pro Asp Leu Asn Arg Phe Leu Thr Arg Thr Leu Gln 820 825 830	2496
GTT CTA CTA CCT GAT CTT GCT GCC AAA GCA AGC CCT GCA GCT TCT GCT Val Leu Leu Pro Asp Leu Ala Ala Lys Ala Ser Pro Ala Ala Ser Ala 835 840 845	2544
CTC ATT CGA ACT TTA GGA AAA CAA TTA AAT GTC AAT CGT AGA GAG ATT Leu Ile Arg Thr Leu Gly Lys Gln Leu Asn Val Asn Arg Arg Glu Ile 850 855 860	2592
TTA ATA AAC AAC TTC AAA TAT ATT TTT TCT CAT TTG GTC TGT TCT TGT Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys 865 870 875 880	2640
TCC AAA GAT GAA TTA GAA CGT GCC CTT CAT TAT CTG AAG AAT GAA ACA Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr 885 890 895	2688
GAA ATT GAA CTG GGG AGC CTG TTG AGA CAA GAT TTC CAA GGA TTG CAT Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His 900 905 910	2736
AAT GAA TTA TTG CTG CGT ATT GGA GAA CAC TAT CAA CAG GTT TTT AAT Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn 915 920 925	2784
GGT TTG TCA ATA CTT GCC TCA TTT GCA TCC AGT GAT GAT CCA TAT CAG Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln 930 935 940	2832
GGC CCG AGA GAT ATC ATA TCA CCT GAA CTG ATG GCT GAT TAT TTA CAA Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln 945 950 955 960	2880
CCC AAA TTG TTG GGC ATT TTG GCT TTT TTT AAC ATG CAG TTA CTG AGC Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser 965 970 975	2928
TCT AGT GTT GGC ATT GAA GAT AAG AAA ATG GCC TTG AAC AGT TTG ATG Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met 980 985 990	2976
TCT TTG ATG AAG TTA ATG GGA CCC AAA CAT GTC AGT TCT GTG AGG GTG Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val 995 1000 1005	3024
AAG ATG ATG ACC ACA CTG AGA ACT GGC CTT CGA TTC AAG GAT GAT TTT Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe 1010 1015 1020	3072
CCT GAA TTG TGT TGC AGA GCT TGG GAC TGC TTT GTT CGC TGC CTG GAT Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp 1025 1030 1035 1040	3120

-169-

CAT GCT TGT CTG GGC TCC CTT CTC AGT CAT GTA ATA GTA GCT TTG TTA His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu 1045 1050 1055	3168
CCT CTT ATA CAC ATC CAG CCT AAA GAA ACT GCA GCT ATC TTC CAC TAC Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr 1060 1065 1070	3216
CTC ATA ATT GAA AAC AGG GAT GCT GTG CAA GAT TTT CTT CAT GAA ATA Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile 1075 1080 1085	3264
TAT TTT TTA CCT GAT CAT CCA GAA TTA AAA AAG ATA AAA GCC GTT CTC Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu 1090 1095 1100	3312
CAG GAA TAC AGA AAG GAG ACC TCT GAG AGC ACT GAT CTT CAG ACA ACT Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr 1105 1110 1115 1120	3360
CTT CAG CTC TCT ATG AAG GCC ATT CAA CAT GAA AAT GTC GAT GTT CGT Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg 1125 1130 1135	3408
ATT CAT GCT CTT ACA AGC TTG AAG GAA ACC TTG TAT AAA AAT CAG GAA Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu 1140 1145 1150	3456
AAA CTG ATA AAG TAT GCA ACA GAC AGT GAA ACA GTA GAA CCT ATT ATC Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile 1155 1160 1165	3504
TCA CAG TTG GTG ACA GTG CTT TTG AAA GGT TGC CAA GAT GCA AAC TCT Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser 1170 1175 1180	3552
CAA GCT CGG TTG CTC TGT GGG GAA TGT TTA GGG GAA TTG GGG GCG ATA Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile 1185 1190 1195 1200	3600
GAT CCA GGT CGA TTA GAT TTC TCA ACA ACT GAA ACT CAA GGA AAA GAT Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp 1205 1210 1215	3648
TTT ACA TTT GTG ACT GGA GTA GAA GAT TCA AGC TTT GCC TAT GGA TTA Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu 1220 1225 1230	3696
TTG ATG GAG CTA ACA AGA GCT TAC CTT GCG TAT GCT GAT AAT AGC CGA Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg 1235 1240 1245	3744
GCT CAA GAT TCA GCT GCC TAT GCC ATT CAG GAG TTG CTT TCT ATT TAT Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr 1250 1255 1260	3792
GAC TGT AGA GAG ATG GAG ACC AAC GGC CCA GGT CAC CAA TTG TGG AGG Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg 1265 1270 1275 1280	3840
AGA TTT CCT GAG CAT GTT CGG GAA ATA CTA GAA CCT CAT CTA AAT ACC Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr 1285 1290 1295	3888
AGA TAC AAG AGT TCT CAG AAG TCA ACC GAT TGG TCT GGA GTA AAG AAG Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys 1300 1305 1310	3936

-170-

CCA ATT TAC TTA AGT AAA TTG GGT AGT AAC TTT GCA GAA TGG TCA GCA Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325	3984
TCT TGG GCA GGT TAT CTT ATT ACA AAG GTT CGA CAT GAT CTT GCC AGT Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1335 1340	4032
AAA ATT TTC ACC TGC TGT AGC ATT ATG ATG AAG CAT GAT TTC AAA GTG Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360	4080
ACC ATC TAT CTT CTT CCA CAT ATT CTG GTG TAT GTC TTA CTG GGT TGT Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375	4128
AAT CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT ATG GCA GTT CTA Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390	4176
AAG CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC ATT GCA TCT GAT Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 1400 1405	4224
CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420	4272
ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440	4320
CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455	4368
GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470	4416
CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485	4464
CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500	4512
ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515 1520	4560
GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535	4608
TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550	4656
GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565	4704
ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1575 1580	4752

-171-

CTG TCT ACT GTT ATC ACT CAG GTG AAT GGA GTG CAT GCT AAC AGG TCC Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600	4800
GAG TGG ACA GAT GAA TTA AAC ACG TAC AGA GTG GAA GCA GCT TGG AAA Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615	4848
TTG TCA CAG TGG GAT TTG GTG GAA AAC TAT TTG GCA GCA GAT GGA AAA Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630	4896
TCT ACA ACA TGG AGT GTC AGA CTG GGA CAG CTA TTA TTA TCA GCC AAA Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys 1635 1640 1645	4944
AAA AGA GAT ATC ACA GCT TTT TAT GAC TCA CTG AAA CTA GTG AGA GCA Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660	4992
GAA CAA ATT GTA CCT CTT TCA GCT GCA AGC TTT GAA AGA GGC TCC TAC Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1665 1670 1675 1680	5040
CAA CGA GGA TAT GAA TAT ATT GTG AGA TTG CAC ATG TTA TGT GAG TTG Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 1695	5088
GAG CAT AGC ATC AAA CCA CTT TTC CAG CAT TCT CCA GGT GAC AGT TCT Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser 1700 1705 1710	5136
CAA GAA GAT TCT CTA AAC TGG GTA GCT CGA CTA GAA ATG ACC CAG AAT Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn 1715 1720 1725	5184
TCC TAC AGA GCC AAG GAG CCT ATC CTG GCT CTC CGG AGG GCT TTA CTA Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1735 1740	5232
AGC CTC AAC AAA AGA CCA GAT TAC AAT GAA ATG GTT GGA GAA TGC TGG Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp 1745 1750 1755 1760	5280
CTG CAG AGT GCC AGG GTA GCT AGA AAG GCT GGT CAC CAC CAG ACA GCC Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala 1765 1770 1775	5328
TAC AAT GCT CTC CTT AAT GCA GGG GAA TCA CGA CTC GCT GAA CTG TAC Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr 1780 1785 1790	5376
GTG GAA AGG GCA AAG TGG CTC TGG TCC AAG GGT GAT GTT CAC CAG GCA Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1795 1800 1805	5424
CTA ATT GTT CTT CAA AAA GGT GTT GAA TTA TGT TTT CCT GAA AAT GAA Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1810 1815 1820	5472
ACC CCA CCT GAG GGT AAG AAC ATG TTA ATC CAT GGT CGA GCT ATG CTA Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu 1825 1830 1835 1840	5520
CTA GTG GGC CGA TTT ATG GAA GAA ACA GCT AAC TTT GAA AGC AAT GCA Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1845 1850 1855	5568

-172-

ATT ATG AAA AAA TAT AAG GAT GTG ACC GCG TGC CTG CCA GAA TGG GAG Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1860 1865 1870	5616
GAT GGG CAT TTT TAC CTT GCC AAG TAC TAT GAC AAA TTG ATG CCC ATG Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1875 1880 1885	5664
GTC ACA GAC AAC AAA ATG GAA AAG CAA GGT GAT CTC ATC CGG TAT ATA Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1890 1895 1900	5712
GTT CTT CAT TTT GGC AGA TCT CTA CAA TAT GGA AAT CAG TTC ATA TAT Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1905 1910 1915 1920	5760
CAG TCA ATG CCA CGA ATG TTA ACT CTA TGG CTT GAT TAT GGT ACA AAG Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 1935	5808
GCA TAT GAA TGG GAA AAA GCT GGC CGC TCC GAT CGT GTA CAA ATG AGG Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1940 1945 1950	5856
AAT GAT TTG GGT AAA ATA AAC AAG GTT ATC ACA GAG CAT ACA AAC TAT Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 1965	5904
TTA GCT CCA TAT CAA TTT TTG ACT GCT TTT TCA CAA TTG ATC TCT CGA Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980	5952
ATT TGT CAT TCT CAC GAT GAA GTT TTT GTT GTC TTG ATG GAA ATA ATA Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Met Glu Ile Ile 1985 1990 1995 2000	6000
GCC AAA GTA TTT CTA GCC TAT CCT CAA CAA GCA ATG TGG ATG ATG ACA Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr 2005 2010 2015	6048
GCT GTG TCA AAG TCA TCT TAT CCC ATG CGT GTG AAC AGA TGC AAG GAA Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030	6096
ATC CTC AAT AAA GCT ATT CAT ATG AAA AAA TCC TTA GAG AAG TTT GTT Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045	6144
GGA GAT GCA ACT CGC CTA ACA GAT AAG CTT CTA GAA TTG TGC AAT AAA Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys 2050 2055 2060	6192
CCG GTT GAT GGA AGT AGT TCC ACA TTA AGC ATG AGC ACT CAT TTT AAA Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys 2065 2070 2075 2080	6240
ATG CTT AAA AAG CTG GTA GAA GAA GCA ACA TTT AGT GAA ATC CTC ATT Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile 2085 2090 2095	6288
CCT CTA CAA TCA GTC ATG ATA CCT ACA CTT CCA TCA ATT CTG GGT ACC Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr 2100 2105 2110	6336
CAT GCT AAC CAT GCT AGC CAT GAA CCA TTT CCT GGA CAT TGG GCC TAT His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr 2115 2120 2125	6384

-173-

ATT GCA GGG TTT GAT GAT ATG GTG GAA ATT CTT GCT TCT CTT CAG AAA Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys 2130 2135 2140	6432
CCA AAG AAG ATT TCT TTA AAA GGC TCA GAT GGA AAG TTC TAC ATC ATG Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met 2145 2150 2155 2160	6480
ATG TGT AAG CCA AAA GAT GAC CTG AGA AAG GAT TGT AGA CTA ATG GAA Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu 2165 2170 2175	6528
TTC AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG TCT CGT Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg 2180 2185 2190	6576
AGA AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp 2195 2200 2205	6624
GAA TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG AGA CCT Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro 2210 2215 2220	6672
ATT CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys 2225 2230 2235 2240	6720
GAA CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys 2245 2250 2255	6768
CTC AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe 2260 2265 2270	6816
CAT GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser 2275 2280 2285	6864
AGT AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly 2290 2295 2300	6912
TAT ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp 2305 2310 2315 2320	6960
TCT TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn 2325 2330 2335	7008
AAG GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr 2340 2345 2350	7056
CAT AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe 2355 2360 2365	7104
CGA AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT CAG CGA GAG Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu 2370 2375 2380	7152
CCT TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT CTT GTG GAA Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu 2385 2390 2395 2400	7200

-174-

TGG AGT AAA CCA GTG AAA GGG CAT TCC AAA GCG CCA CTG AAT GAA ACT	7248
Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr	
2405 2410 2415	
GGA GAA GTT GTC AAT GAA AAG GCC AAG ACC CAT GTT CTT GAC ATT GAG	7296
Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu	
2420 2425 2430	
CAG CGA CTA CAA GGT GTA ATC AAG ACT CGA AAT AGA GTG ACA GGA CTG	7344
Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu	
2435 2440 2445	
CCG TTA TCT ATT GAA GGA CAT GTG CAT TAC CTT ATA CAA GAA GCT ACT	7392
Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr	
2450 2455 2460	
GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT CCA TAT ATG	7440
Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met	
2465 2470 2475 2480	
TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA AGTAAAAAAA AAAAAAAAAA	7500
AA	7502

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2480 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser	
1 5 10 15	
Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu	
20 25 30	
Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val	
35 40 45	
Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu	
50 55 60	
Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile	
65 70 75 80	
Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly	
85 90 95	
Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu	
100 105 110	
Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu	
115 120 125	
Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu	
130 135 140	
Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys	
145 150 155 160	

-175-

Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp
 165 170 175
 Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe
 180 185 190
 Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val
 195 200 205
 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp
 210 215 220
 Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser
 225 230 235 240
 Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu
 245 250 255
 Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser
 260 265 270
 Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys
 275 280 285
 His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln
 290 295 300
 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn
 305 310 315 320
 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr
 325 330 335
 Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr
 340 345 350
 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile
 355 360 365
 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg
 370 375 380
 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp
 385 390 395 400
 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser
 405 410 415
 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu
 420 425 430
 Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr
 435 440 445
 Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr
 450 455 460
 Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro
 465 470 475 480
 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu
 485 490 495
 Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe
 500 505 510

-176-

Ile	Leu	Leu	Gln	Gln	Gln	Asn	Ser	Cys	Asn	Arg	Val	Pro	Lys	Ile	Leu
	515						520					525			
Ile	Asp	Lys	Val	Lys	Asp	Asp	Ser	Asp	Ile	Val	Lys	Lys	Glu	Phe	Ala
	530					535					540				
Ser	Ile	Leu	Gly	Gln	Leu	Val	Cys	Thr	Leu	His	Gly	Met	Phe	Tyr	Leu
545					550					555					560
Thr	Ser	Ser	Leu	Thr	Glu	Pro	Phe	Ser	Glu	His	Gly	His	Val	Asp	Leu
				565					570					575	
Phe	Cys	Arg	Asn	Leu	Lys	Ala	Thr	Ser	Gln	His	Glu	Cys	Ser	Ser	Ser
			580					585					590		
Gln	Leu	Lys	Ala	Ser	Val	Cys	Lys	Pro	Phe	Leu	Phe	Leu	Leu	Lys	Lys
		595					600					605			
Lys	Ile	Pro	Ser	Pro	Val	Lys	Leu	Ala	Phe	Ile	Asp	Asn	Leu	His	His
	610					615					620				
Leu	Cys	Lys	His	Leu	Asp	Phe	Arg	Glu	Asp	Glu	Thr	Asp	Val	Lys	Ala
625					630					635					640
Val	Leu	Gly	Thr	Leu	Leu	Asn	Leu	Met	Glu	Asp	Pro	Asp	Lys	Asp	Val
				645					650					655	
Arg	Val	Ala	Phe	Ser	Gly	Asn	Ile	Lys	His	Ile	Leu	Glu	Ser	Leu	Asp
			660					665					670		
Ser	Glu	Asp	Gly	Phe	Ile	Lys	Glu	Leu	Phe	Val	Leu	Arg	Met	Lys	Glu
		675					680					685			
Ala	Tyr	Thr	His	Ala	Gln	Ile	Ser	Arg	Asn	Asn	Glu	Leu	Lys	Asp	Thr
	690					695					700				
Leu	Ile	Leu	Thr	Thr	Gly	Asp	Ile	Gly	Arg	Ala	Ala	Lys	Gly	Asp	Leu
705					710					715					720
Val	Pro	Phe	Ala	Leu	Leu	His	Leu	Leu	His	Cys	Leu	Leu	Ser	Lys	Ser
				725					730					735	
Ala	Ser	Val	Ser	Gly	Ala	Ala	Tyr	Thr	Glu	Ile	Arg	Ala	Leu	Val	Ala
			740					745					750		
Ala	Lys	Ser	Val	Lys	Leu	Gln	Ser	Phe	Phe	Ser	Gln	Tyr	Lys	Lys	Pro
		755					760					765			
Ile	Cys	Gln	Phe	Leu	Val	Glu	Ser	Leu	His	Ser	Ser	Gln	Met	Thr	Ala
	770					775						780			
Leu	Pro	Asn	Thr	Pro	Cys	Gln	Asn	Ala	Asp	Val	Arg	Lys	Gln	Asp	Val
785					790					795					800
Ala	His	Gln	Arg	Glu	Met	Ala	Leu	Asn	Thr	Leu	Ser	Glu	Ile	Ala	Asn
				805					810					815	
Val	Phe	Asp	Phe	Pro	Asp	Leu	Asn	Arg	Phe	Leu	Thr	Arg	Thr	Leu	Gln
			820					825					830		
Val	Leu	Leu	Pro	Asp	Leu	Ala	Ala	Lys	Ala	Ser	Pro	Ala	Ala	Ser	Ala
		835					840					845			
Leu	Ile	Arg	Thr	Leu	Gly	Lys	Gln	Leu	Asn	Val	Asn	Arg	Arg	Glu	Ile
	850					855					860				

-177-

Leu Ile Asn Asn Phe Lys Tyr Ile Phe Ser His Leu Val Cys Ser Cys
 865 870 875 880
 Ser Lys Asp Glu Leu Glu Arg Ala Leu His Tyr Leu Lys Asn Glu Thr
 885 890 895
 Glu Ile Glu Leu Gly Ser Leu Leu Arg Gln Asp Phe Gln Gly Leu His
 900 905 910
 Asn Glu Leu Leu Leu Arg Ile Gly Glu His Tyr Gln Gln Val Phe Asn
 915 920 925
 Gly Leu Ser Ile Leu Ala Ser Phe Ala Ser Ser Asp Asp Pro Tyr Gln
 930 935 940
 Gly Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln
 945 950 955 960
 Pro Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser
 965 970 975
 Ser Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met
 980 985 990
 Ser Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val
 995 1000 1005
 Lys Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe
 1010 1015 1020
 Pro Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp
 1025 1030 1035 1040
 His Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu
 1045 1050 1055
 Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr
 1060 1065 1070
 Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile
 1075 1080 1085
 Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu
 1090 1095 1100
 Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr
 1105 1110 1115 1120
 Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg
 1125 1130 1135
 Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu
 1140 1145 1150
 Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile
 1155 1160 1165
 Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser
 1170 1175 1180
 Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile
 1185 1190 1195 1200
 Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp
 1205 1210 1215

-178-

Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu
 1220 1225 1230
 Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg
 1235 1240 1245
 Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr
 1250 1255 1260
 Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg
 1265 1270 1275 1280
 Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr
 1285 1290 1295
 Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys
 1300 1305 1310
 Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala
 1315 1320 1325
 Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser
 1330 1335 1340
 Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val
 1345 1350 1355 1360
 Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys
 1365 1370 1375
 Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu
 1380 1385 1390
 Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp
 1395 1400 1405
 Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu
 1410 1415 1420
 Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys
 1425 1430 1435 1440
 Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr
 1445 1450 1455
 Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile
 1460 1465 1470
 Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr
 1475 1480 1485
 Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn
 1490 1495 1500
 Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His
 1505 1510 1515 1520
 Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro
 1525 1530 1535
 Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg
 1540 1545 1550
 Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln
 1555 1560 1565

-179-

Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln
 1570 1575 1580
 Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser
 1585 1590 1595 1600
 Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys
 1605 1610 1615
 Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys
 1620 1625 1630
 Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys
 1635 1640 1645
 Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala
 1650 1655 1660
 Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr
 1665 1670 1675 1680
 Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu
 1685 1690 1695
 Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser
 1700 1705 1710
 Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn
 1715 1720 1725
 Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu
 1730 1735 1740
 Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp
 1745 1750 1755 1760
 Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala
 1765 1770 1775
 Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr
 1780 1785 1790
 Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala
 1795 1800 1805
 Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu
 1810 1815 1820
 Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu
 1825 1830 1835 1840
 Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala
 1845 1850 1855
 Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu
 1860 1865 1870
 Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met
 1875 1880 1885
 Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile
 1890 1895 1900
 Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr
 1905 1910 1915 1920

-180-

Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys
 1925 1930 1935
 Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg
 1940 1945 1950
 Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr
 1955 1960 1965
 Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg
 1970 1975 1980
 Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Met Glu Ile Ile
 1985 1990 1995 2000
 Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr
 2005 2010 2015
 Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu
 2020 2025 2030
 Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val
 2035 2040 2045
 Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys
 2050 2055 2060
 Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys
 2065 2070 2075 2080
 Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile
 2085 2090 2095
 Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr
 2100 2105 2110
 His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr
 2115 2120 2125
 Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys
 2130 2135 2140
 Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met
 2145 2150 2155 2160
 Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu
 2165 2170 2175
 Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg
 2180 2185 2190
 Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp
 2195 2200 2205
 Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro
 2210 2215 2220
 Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys
 2225 2230 2235 2240
 Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys
 2245 2250 2255
 Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe
 2260 2265 2270

-181-

His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser
 2275 2280 2285
 Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly
 2290 2295 2300
 Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp
 2305 2310 2315 2320
 Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn
 2325 2330 2335
 Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr
 2340 2345 2350
 His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe
 2355 2360 2365
 Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu
 2370 2375 2380
 Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu
 2385 2390 2395 2400
 Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr
 2405 2410 2415
 Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu
 2420 2425 2430
 Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu
 2435 2440 2445
 Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr
 2450 2455 2460
 Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met
 2465 2470 2475 2480

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 9385 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(v) FRAGMENT TYPE: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 190..9357

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GCGAGAGGAG TCGGGATCTG CGCTGCAGCC ACCGCCGCGG TTGATACTAC TTTGACCTTC	60
CGAGTGCACT GAGGCATACA TCACAATTTG GAATTATGCA TTGGTTTATC AATTTACTTG	120
TTTATTGTCA CCCTGCTGCC CAGATATGAC TTCATGAGGA CAGTGATGTG TGTCTGAAA	180

-182-

TTGTGAACC	ATG	AGT	CTA	GTA	CTT	AAT	GAT	CTG	CTT	ATC	TGC	TGC	CGT	228
	Met	Ser	Leu	Val	Leu	Asn	Asp	Leu	Leu	Ile	Cys	Cys	Arg	
	1				5					10				
CAA	CTA	GAA	CAT	GAT	AGA	GCT	ACA	GAA	CGA	AAG	AAA	GAA	GTT	276
Gln	Leu	Glu	His	Asp	Arg	Ala	Thr	Glu	Arg	Lys	Lys	Glu	Val	
	15					20					25		Glu	
TTT	AAG	CGC	CTG	ATT	CGA	GAT	CCT	GAA	ACA	ATT	AAA	CAT	CTA	324
Phe	Lys	Arg	Leu	Ile	Arg	Asp	Pro	Glu	Thr	Ile	Lys	His	Leu	
	30				35					40			Asp	
													Arg	45
CAT	TCA	GAT	TCC	AAA	CAA	GGA	AAA	TAT	TTG	AAT	TGG	GAT	GCT	372
His	Ser	Asp	Ser	Lys	Gln	Gly	Lys	Tyr	Leu	Asn	Trp	Asp	Ala	
				50					55				Val	
													60	
													Phe	
AGA	TTT	TTA	CAG	AAA	TAT	ATT	CAG	AAA	GAA	ACA	GAA	TGT	CTG	420
Arg	Phe	Leu	Gln	Lys	Tyr	Ile	Gln	Lys	Glu	Thr	Glu	Cys	Leu	
			65					70					Arg	
													Ile	
GCA	AAA	CCA	AAT	GTA	TCA	GCC	TCA	ACA	CAA	GCC	TCC	AGG	CAG	468
Ala	Lys	Pro	Asn	Val	Ser	Ala	Ser	Thr	Gln	Ala	Ser	Arg	Gln	
		80					85					90	Lys	
													Lys	
ATG	CAG	GAA	ATC	AGT	AGT	TTG	GTC	AAA	TAC	TTC	ATC	AAA	TGT	516
Met	Gln	Glu	Ile	Ser	Ser	Leu	Val	Lys	Tyr	Phe	Ile	Lys	Cys	
	95					100					105		Ala	
													Asn	
AGA	AGA	GCA	CCT	AGG	CTA	AAA	TGT	CAA	GAA	CTC	TTA	AAT	TAT	564
Arg	Arg	Ala	Pro	Arg	Leu	Lys	Cys	Gln	Glu	Leu	Leu	Asn	Tyr	
110					115					120			Ile	
													Met	
GAT	ACA	GTG	AAA	GAT	TCA	TCT	AAT	GGT	GCT	ATT	TAC	GGA	GCT	612
Asp	Thr	Val	Lys	Asp	Ser	Ser	Asn	Gly	Ala	Ile	Tyr	Gly	Ala	
				130					135				Asp	
													Cys	
AGC	AAC	ATA	CTA	CTC	AAA	GAC	ATT	CTT	TCT	GTG	AGA	AAA	TAC	660
Ser	Asn	Ile	Leu	Leu	Lys	Asp	Ile	Leu	Ser	Val	Arg	Lys	Tyr	
			145					150					Trp	
													Cys	
GAA	ATA	TCT	CAG	CAA	CAG	TGG	TTA	GAA	TTG	TTC	TCT	GTG	TAC	708
Glu	Ile	Ser	Gln	Gln	Gln	Trp	Leu	Glu	Leu	Phe	Ser	Val	Tyr	
			160				165					170	Phe	
													Arg	
CTC	TAT	CTG	AAA	CCT	TCA	CAA	GAT	GTT	CAT	AGA	GTT	TTA	GTG	756
Leu	Tyr	Leu	Lys	Pro	Ser	Gln	Asp	Val	His	Arg	Val	Leu	Val	
						180					185		Ala	
													Arg	
ATA	ATT	CAT	GCT	GTT	ACC	AAA	GGA	TGC	TGT	TCT	CAG	ACT	GAC	804
Ile	Ile	His	Ala	Val	Thr	Lys	Gly	Cys	Cys	Ser	Gln	Thr	Asp	
190					195					200			Gly	
													Leu	
AAT	TCC	AAA	TTT	TTG	GAC	TTT	TTT	TCC	AAG	GCT	ATT	CAG	TGT	852
Asn	Ser	Lys	Phe	Leu	Asp	Phe	Phe	Ser	Lys	Ala	Ile	Gln	Cys	
				210					215				Ala	
													Arg	
CAA	GAA	AAG	AGC	TCT	TCA	GGT	CTA	AAT	CAT	ATC	TTA	GCA	GCT	900
Gln	Glu	Lys	Ser	Ser	Ser	Gly	Leu	Asn	His	Ile	Leu	Ala	Ala	
			225				230					235	Leu	
													Thr	
ATC	TTC	CTC	AAG	ACT	TTG	GCT	GTC	AAC	TTT	CGA	ATT	CGA	GTG	948
Ile	Phe	Leu	Lys	Thr	Leu	Ala	Val	Asn	Phe	Arg	Ile	Arg	Val	
			240				245					250	Cys	
													Glu	
TTA	GGA	GAT	GAA	ATT	CTT	CCC	ACT	TTG	CTT	TAT	ATT	TGG	ACT	996
Leu	Gly	Asp	Glu	Ile	Leu	Pro	Thr	Leu	Leu	Tyr	Ile	Trp	Thr	
	255					260					265		Gln	
													His	

-183-

AGG CTT AAT GAT TCT TTA AAA GAA GTC ATT ATT GAA TTA TTT CAA CTG Arg Leu Asn Asp Ser Leu Lys Glu Val Ile Ile Glu Leu Phe Gln Leu 270 275 280 285	1044
CAA ATT TAT ATC CAT CAT CCG AAA GGA GCC AAA ACC CAA GAA AAA GGT Gln Ile Tyr Ile His His Pro Lys Gly Ala Lys Thr Gln Glu Lys Gly 290 295 300	1092
GCT TAT GAA TCA ACA AAA TGG AGA AGT ATT TTA TAC AAC TTA TAT GAT Ala Tyr Glu Ser Thr Lys Trp Arg Ser Ile Leu Tyr Asn Leu Tyr Asp 305 310 315	1140
CTG CTA GTG AAT GAG ATA AGT CAT ATA GGA AGT AGA GGA AAG TAT TCT Leu Leu Val Asn Glu Ile Ser His Ile Gly Ser Arg Gly Lys Tyr Ser 320 325 330	1188
TCA GGA TTT CGT AAT ATT GCC GTC AAA GAA AAT TTG ATT GAA TTG ATG Ser Gly Phe Arg Asn Ile Ala Val Lys Glu Asn Leu Ile Glu Leu Met 335 340 345	1236
GCA GAT ATC TGT CAC CAG GTT TTT AAT GAA GAT ACC AGA TCC TTG GAG Ala Asp Ile Cys His Gln Val Phe Asn Glu Asp Thr Arg Ser Leu Glu 350 355 360 365	1284
ATT TCT CAA TCT TAC ACT ACT ACA CAA AGA GAA TCT AGT GAT TAC AGT Ile Ser Gln Ser Tyr Thr Thr Thr Gln Arg Glu Ser Ser Asp Tyr Ser 370 375 380	1332
GTC CCT TGC AAA AGG AAG AAA ATA GAA CTA GGC TGG GAA GTA ATA AAA Val Pro Cys Lys Arg Lys Lys Ile Glu Leu Gly Trp Glu Val Ile Lys 385 390 395	1380
GAT CAC CTT CAG AAG TCA CAG AAT GAT TTT GAT CTT GTG CCT TGG CTA Asp His Leu Gln Lys Ser Gln Asn Asp Phe Asp Leu Val Pro Trp Leu 400 405 410	1428
CAG ATT GCA ACC CAA TTA ATA TCA AAG TAT CCT GCA AGT TTA CCT AAC Gln Ile Ala Thr Gln Leu Ile Ser Lys Tyr Pro Ala Ser Leu Pro Asn 415 420 425	1476
TGT GAG CTG TCT CCA TTA CTG ATG ATA CTA TCT CAG CTT CTA CCC CAA Cys Glu Leu Ser Pro Leu Leu Met Ile Leu Ser Gln Leu Leu Pro Gln 430 435 440 445	1524
CAG CGA CAT GGG GAA CGT ACA CCA TAT GTG TTA CGA TGC CTT ACG GAA Gln Arg His Gly Glu Arg Thr Pro Tyr Val Leu Arg Cys Leu Thr Glu 450 455 460	1572
GTT GCA TTG TGT CAA GAC AAG AGG TCA AAC CTA GAA AGC TCA CAA AAG Val Ala Leu Cys Gln Asp Lys Arg Ser Asn Leu Glu Ser Ser Gln Lys 465 470 475	1620
TCA GAT TTA TTA AAA CTC TGG AAT AAA ATT TGG TGT ATT ACC TTT CGT Ser Asp Leu Leu Lys Leu Trp Asn Lys Ile Trp Cys Ile Thr Phe Arg 480 485 490	1668
GGT ATA AGT TCT GAG CAA ATA CAA GCT GAA AAC TTT GGC TTA CTT GGA Gly Ile Ser Ser Glu Gln Ile Gln Ala Glu Asn Phe Gly Leu Leu Gly 495 500 505	1716
GCC ATA ATT CAG GGT AGT TTA GTT GAG GTT GAC AGA GAA TTC TGG AAG Ala Ile Ile Gln Gly Ser Leu Val Glu Val Asp Arg Glu Phe Trp Lys 510 515 520 525	1764
TTA TTT ACT GGG TCA GCC TGC AGA CCT TCA TGT CCT GCA GTA TGC TGT Leu Phe Thr Gly Ser Ala Cys Arg Pro Ser Cys Pro Ala Val Cys Cys 530 535 540	1812

-184-

TTG ACT TTG GCA CTG ACC ACC AGT ATA GTT CCA GGA GCG GTA AAA ATG Leu Thr Leu Ala Leu Thr Thr Ser Ile Val Pro Gly Ala Val Lys Met 545 550 555	1860
GGA ATA GAG CAA AAT ATG TGT GAA GTA AAT AGA AGC TTT TCT TTA AAG Gly Ile Glu Gln Asn Met Cys Glu Val Asn Arg Ser Phe Ser Leu Lys 560 565 570	1908
GAA TCA ATA ATG AAA TGG CTC TTA TTC TAT CAG TTA GAG GGT GAC TTA Glu Ser Ile Met Lys Trp Leu Leu Phe Tyr Gln Leu Glu Gly Asp Leu 575 580 585	1956
GAA AAT AGC ACA GAA GTG CCT CCA ATT CTT CAC AGT AAT TTT CCT CAT Glu Asn Ser Thr Glu Val Pro Pro Ile Leu His Ser Asn Phe Pro His 590 595 600 605	2004
CTT GTA CTG GAG AAA ATT CTT GTG AGT CTC ACT ATG AAA AAC TGT AAA Leu Val Leu Glu Lys Ile Leu Val Ser Leu Thr Met Lys Asn Cys Lys 610 615 620	2052
GCT GCA ATG AAT TTT TTC CAA AGC GTG CCA GAA TGT GAA CAC CAC CAA Ala Ala Met Asn Phe Phe Gln Ser Val Pro Glu Cys Glu His His Gln 625 630 635	2100
AAA GAT AAA GAA GAA CTT TCA TTC TCA GAA GTA GAA GAA CTA TTT CTT Lys Asp Lys Glu Glu Leu Ser Phe Ser Glu Val Glu Glu Leu Phe Leu 640 645 650	2148
CAG ACA ACT TTT GAC AAG ATG GAC TTT TTA ACC ATT GTG AGA GAA TGT Gln Thr Thr Phe Asp Lys Met Asp Phe Leu Thr Ile Val Arg Glu Cys 655 660 665	2196
GGT ATA GAA AAG CAC CAG TCC AGT ATT GGC TTC TCT GTC CAC CAG AAT Gly Ile Glu Lys His Gln Ser Ser Ile Gly Phe Ser Val His Gln Asn 670 675 680 685	2244
CTC AAG GAA TCA CTG GAT CGC TGT CTT CTG GGA TTA TCA GAA CAG CTT Leu Lys Glu Ser Leu Asp Arg Cys Leu Leu Gly Leu Ser Glu Gln Leu 690 695 700	2292
CTG AAT AAT TAC TCA TCT GAG ATT ACA AAT TCA GAA ACT CTT GTC CGG Leu Asn Asn Tyr Ser Ser Glu Ile Thr Asn Ser Glu Thr Leu Val Arg 705 710 715	2340
TGT TCA CGT CTT TTG GTG GGT GTC CTT GGC TGC TAC TGT TAC ATG GGT Cys Ser Arg Leu Leu Val Gly Val Leu Gly Cys Tyr Cys Tyr Met Gly 720 725 730	2388
GTA ATA GCT GAA GAG GAA GCA TAT AAG TCA GAA TTA TTC CAG AAA GCC Val Ile Ala Glu Glu Glu Ala Tyr Lys Ser Glu Leu Phe Gln Lys Ala 735 740 745	2436
AAC TCT CTA ATG CAA TGT GCA GGA GAA AGT ATC ACT CTG TTT AAA AAT Asn Ser Leu Met Gln Cys Ala Gly Glu Ser Ile Thr Leu Phe Lys Asn 750 755 760 765	2484
AAG ACA AAT GAG GAA TTC AGA ATT GGT TCC TTG AGA AAT ATG ATG CAG Lys Thr Asn Glu Phe Arg Ile Gly Ser Leu Arg Asn Met Met Gln 770 775 780	2532
CTA TGT ACA CGT TGC TTG AGC AAC TGT ACC AAG AAG AGT CCA AAT AAG Leu Cys Thr Arg Cys Leu Ser Asn Cys Thr Lys Lys Ser Pro Asn Lys 785 790 795	2580
ATT GCA TCT GGC TTT TTC CTG CGA TTG TTA ACA TCA AAG CTA ATG AAT Ile Ala Ser Gly Phe Phe Leu Arg Leu Leu Thr Ser Lys Leu Met Asn 800 805 810	2628

-185-

GAC ATT GCA GAT ATT TGT AAA AGT TTA GCA TCC TTC ATC AAA AAG CCA Asp Ile Ala Asp Ile Cys Lys Ser Leu Ala Ser Phe Ile Lys Lys Pro 815 820 825	2676
TTT GAC CGT GGA GAA GTA GAA TCA ATG GAA GAT GAT ACT AAT GGA AAT Phe Asp Arg Gly Glu Val Glu Ser Met Glu Asp Asp Thr Asn Gly Asn 830 835 840 845	2724
CTA ATG GAG GTG GAG GAT CAG TCA TCC ATG AAT CTA TTT AAC GAT TAC Leu Met Glu Val Glu Asp Gln Ser Ser Met Asn Leu Phe Asn Asp Tyr 850 855 860	2772
CCT GAT AGT AGT GTT AGT GAT GCA AAC GAA CCT GGA GAG AGC CAA AGT Pro Asp Ser Ser Val Ser Asp Ala Asn Glu Pro Gly Glu Ser Gln Ser 865 870 875	2820
ACC ATA GGT GCC ATT AAT CCT TTA GCT GAA GAA TAT CTG TCA AAG CAA Thr Ile Gly Ala Ile Asn Pro Leu Ala Glu Glu Tyr Leu Ser Lys Gln 880 885 890	2868
GAT CTA CTT TTC TTA GAC ATG CTC AAG TTC TTG TGT TTG TGT GTA ACT Asp Leu Leu Phe Leu Asp Met Leu Lys Phe Leu Cys Leu Cys Val Thr 895 900 905	2916
ACT GCT CAG ACC AAT ACT GTG TCC TTT AGG GCA GCT GAT ATT CGG AGG Thr Ala Gln Thr Asn Thr Val Ser Phe Arg Ala Ala Asp Ile Arg Arg 910 915 920 925	2964
AAA TTG TTA ATG TTA ATT GAT TCT AGC ACG CTA GAA CCT ACC AAA TCC Lys Leu Leu Met Leu Ile Asp Ser Ser Thr Leu Glu Pro Thr Lys Ser 930 935 940	3012
CTC CAC CTG CAT ATG TAT CTA ATG CTT TTA AAG GAG CTT CCT GGA GAA Leu His Leu His Met Tyr Leu Met Leu Leu Lys Glu Leu Pro Gly Glu 945 950 955	3060
GAG TAC CCC TTG CCA ATG GAA GAT GTT CTT GAA CTT CTG AAA CCA CTA Glu Tyr Pro Leu Pro Met Glu Asp Val Leu Glu Leu Leu Lys Pro Leu 960 965 970	3108
TCC AAT GTG TGT TCT TTG TAT CGT CGT GAC CAA GAT GTT TGT AAA ACT Ser Asn Val Cys Ser Leu Tyr Arg Arg Asp Gln Asp Val Cys Lys Thr 975 980 985	3156
ATT TTA AAC CAT GTC CTT CAT GTA GTG AAA AAC CTA GGT CAA AGC AAT Ile Leu Asn His Val Leu His Val Val Lys Asn Leu Gly Gln Ser Asn 990 995 1000 1005	3204
ATG GAC TCT GAG AAC ACA AGG GAT GCT CAA GGA CAG TTT CTT ACA GTA Met Asp Ser Glu Asn Thr Arg Asp Ala Gln Gly Gln Phe Leu Thr Val 1010 1015 1020	3252
ATT GGA GCA TTT TGG CAT CTA ACA AAG GAG AGG AAA TAT ATA TTC TCT Ile Gly Ala Phe Trp His Leu Thr Lys Glu Arg Lys Tyr Ile Phe Ser 1025 1030 1035	3300
GTA AGA ATG GCC CTA GTA AAT TGC CTT AAA ACT TTG CTT GAG GCT GAT Val Arg Met Ala Leu Val Asn Cys Leu Lys Thr Leu Leu Glu Ala Asp 1040 1045 1050	3348
CCT TAT TCA AAA TGG GCC ATT CTT AAT GTA ATG GGA AAA GAC TTT CCT Pro Tyr Ser Lys Trp Ala Ile Leu Asn Val Met Gly Lys Asp Phe Pro 1055 1060 1065	3396
GTA AAT GAA GTA TTT ACA CAA TTT CTT GCT GAC AAT CAT CAC CAA GTT Val Asn Glu Val Phe Thr Gln Phe Leu Ala Asp Asn His His Gln Val 1070 1075 1080 1085	3444

-186-

CGC ATG TTG GCT GCA GAG TCA ATC AAT AGA TTG TTC CAG GAC ACG AAG Arg Met Leu Ala Ala Glu Ser Ile Asn Arg Leu Phe Gln Asp Thr Lys 1090 1095 1100	3492
GGA GAT TCT TCC AGG TTA CTG AAA GCA CTT CCT TTG AAG CTT CAG CAA Gly Asp Ser Ser Arg Leu Leu Lys Ala Leu Pro Leu Lys Leu Gln Gln 1105 1110 1115	3540
ACA GCT TTT GAA AAT GCA TAC TTG AAA GCT CAG GAA GGA ATG AGA GAA Thr Ala Phe Glu Asn Ala Tyr Leu Lys Ala Gln Glu Gly Met Arg Glu 1120 1125 1130	3588
ATG TCC CAT AGT GCT GAG AAC CCT GAA ACT TTG GAT GAA ATT TAT AAT Met Ser His Ser Ala Glu Asn Pro Glu Thr Leu Asp Glu Ile Tyr Asn 1135 1140 1145	3636
AGA AAA TCT GTT TTA CTG ACG TTG ATA GCT GTG GTT TTA TCC TGT AGC Arg Lys Ser Val Leu Leu Thr Leu Ile Ala Val Val Leu Ser Cys Ser 1150 1155 1160 1165	3684
CCT ATC TGC GAA AAA CAG GCT TTG TTT GCC CTG TGT AAA TCT GTG AAA Pro Ile Cys Glu Lys Gln Ala Leu Phe Ala Leu Cys Lys Ser Val Lys 1170 1175 1180	3732
GAG AAT GGA TTA GAA CCT CAC CTT GTG AAA AAG GTT TTA GAG AAA GTT Glu Asn Gly Leu Glu Pro His Leu Val Lys Lys Val Leu Glu Lys Val 1185 1190 1195	3780
TCT GAA ACT TTT GGA TAT AGA CGT TTA GAA GAC TTT ATG GCA TCT CAT Ser Glu Thr Phe Gly Tyr Arg Arg Leu Glu Asp Phe Met Ala Ser His 1200 1205 1210	3828
TTA GAT TAT CTG GTT TTG GAA TGG CTA AAT CTT CAA GAT ACT GAA TAC Leu Asp Tyr Leu Val Leu Glu Trp Leu Asn Leu Gln Asp Thr Glu Tyr 1215 1220 1225	3876
AAC TTA TCT TCT TTT CCT TTT ATT TTA TTA AAC TAC ACA AAT ATT GAG Asn Leu Ser Ser Phe Pro Phe Ile Leu Leu Asn Tyr Thr Asn Ile Glu 1230 1235 1240 1245	3924
GAT TTC TAT AGA TCT TGT TAT AAG GTT TTG ATT CCA CAT CTG GTG ATT Asp Phe Tyr Arg Ser Cys Tyr Lys Val Leu Ile Pro His Leu Val Ile 1250 1255 1260	3972
AGA AGT CAT TTT GAT GAG GTG AAG TCC ATT GCT AAT CAG ATT CAA GAG Arg Ser His Phe Asp Glu Val Lys Ser Ile Ala Asn Gln Ile Gln Glu 1265 1270 1275	4020
GAC TGG AAA AGT CTT CTA ACA GAC TGC TTT CCA AAG ATT CTT GTA AAT Asp Trp Lys Ser Leu Leu Thr Asp Cys Phe Pro Lys Ile Leu Val Asn 1280 1285 1290	4068
ATT CTT CCT TAT TTT GCC TAT GAG GGT ACC AGA GAC AGT GGG ATG GCA Ile Leu Pro Tyr Phe Ala Tyr Glu Gly Thr Arg Asp Ser Gly Met Ala 1295 1300 1305	4116
CAG CAA AGA GAG ACT GCT ACC AAG GTC TAT GAT ATG CTT AAA AGT GAA Gln Gln Arg Glu Thr Ala Thr Lys Val Tyr Asp Met Leu Lys Ser Glu 1310 1315 1320 1325	4164
AAC TTA TTG GGA AAA CAG ATT GAT CAC TTA TTC ATT AGT AAT TTA CCA Asn Leu Leu Gly Lys Gln Ile Asp His Leu Phe Ile Ser Asn Leu Pro 1330 1335 1340	4212
GAG ATT GTG GTG GAG TTA TTG ATG ACG TTA CAT GAG CCA GCA AAT TCT Glu Ile Val Val Glu Leu Leu Met Thr Leu His Glu Pro Ala Asn Ser 1345 1350 1355	4260

-187-

AGT GCC AGT CAG AGC ACT GAC CTC TGT GAC TTT TCA GGG GAT TTG GAT Ser Ala Ser Gln Ser Thr Asp Leu Cys Asp Phe Ser Gly Asp Leu Asp 1360 1365 1370	4308
CCT GCT CCT AAT CCA CCT CAT TTT CCA TCG CAT GTG ATT AAA GCA ACA Pro Ala Pro Asn Pro Pro His Phe Pro Ser His Val Ile Lys Ala Thr 1375 1380 1385	4356
TTT GCC TAT ATC AGC AAT TGT CAT AAA ACC AAG TTA AAA AGC ATT TTA Phe Ala Tyr Ile Ser Asn Cys His Lys Thr Lys Leu Lys Ser Ile Leu 1390 1395 1400 1405	4404
GAA ATT CTT TCC AAA AGC CCT GAT TCC TAT CAG AAA ATT CTT CTT GCC Glu Ile Leu Ser Lys Ser Pro Asp Ser Tyr Gln Lys Ile Leu Leu Ala 1410 1415 1420	4452
ATA TGT GAG CAA GCA GCT GAA ACA AAT AAT GTT TAT AAG AAG CAC AGA Ile Cys Glu Gln Ala Ala Glu Thr Asn Asn Val Tyr Lys Lys His Arg 1425 1430 1435	4500
ATT CTT AAA ATA TAT CAC CTG TTT GTT AGT TTA TTA CTG AAA GAT ATA Ile Leu Lys Ile Tyr His Leu Phe Val Ser Leu Leu Leu Lys Asp Ile 1440 1445 1450	4548
AAA AGT GGC TTA GGA GGA GCT TGG GCC TTT GTT CTT CGA GAC GTT ATT Lys Ser Gly Leu Gly Gly Ala Trp Ala Phe Val Leu Arg Asp Val Ile 1455 1460 1465	4596
TAT ACT TTG ATT CAC TAT ATC AAC CAA AGG CCT TCT TGT ATC ATG GAT Tyr Thr Leu Ile His Tyr Ile Asn Gln Arg Pro Ser Cys Ile Met Asp 1470 1475 1480 1485	4644
GTG TCA TTA CGT AGC TTC TCC CTT TGT TGT GAC TTA TTA AGT CAG GTT Val Ser Leu Arg Ser Phe Ser Leu Cys Cys Asp Leu Leu Ser Gln Val 1490 1495 1500	4692
TGC CAG ACA GCC GTG ACT TAC TGT AAG GAT GCT CTA GAA AAC CAT CTT Cys Gln Thr Ala Val Thr Tyr Cys Lys Asp Ala Leu Glu Asn His Leu 1505 1510 1515	4740
CAT GTT ATT GTT GGT ACA CTT ATA CCC CTT GTG TAT GAG CAG GTG GAG His Val Ile Val Gly Thr Leu Ile Pro Leu Val Tyr Glu Gln Val Glu 1520 1525 1530	4788
GTT CAG AAA CAG GTA TTG GAC TTG TTG AAA TAC TTA GTG ATA GAT AAC Val Gln Lys Gln Val Leu Asp Leu Leu Lys Tyr Leu Val Ile Asp Asn 1535 1540 1545	4836
AAG GAT AAT GAA AAC CTC TAT ATC ACG ATT AAG CTT TTA GAT CCT TTT Lys Asp Asn Glu Asn Leu Tyr Ile Thr Ile Lys Leu Leu Asp Pro Phe 1550 1555 1560 1565	4884
CCT GAC CAT GTT GTT TTT AAG GAT TTG CGT ATT ACT CAG CAA AAA ATC Pro Asp His Val Val Phe Lys Asp Leu Arg Ile Thr Gln Gln Lys Ile 1570 1575 1580	4932
AAA TAC AGT AGA GGA CCC TTT TCA CTC TTG GAG GAA ATT AAC CAT TTT Lys Tyr Ser Arg Gly Pro Phe Ser Leu Leu Glu Glu Ile Asn His Phe 1585 1590 1595	4980
CTC TCA GTA AGT GTT TAT GAT GCA CTT CCA TTG ACA AGA CTT GAA GGA Leu Ser Val Ser Val Tyr Asp Ala Leu Pro Leu Thr Arg Leu Glu Gly 1600 1605 1610	5028
CTA AAG GAT CTT CGA AGA CAA CTG GAA CTA CAT AAA GAT CAG ATG GTG Leu Lys Asp Leu Arg Arg Gln Leu Glu Leu His Lys Asp Gln Met Val 1615 1620 1625	5076

-188-

GAC ATT ATG AGA GCT TCT CAG GAT AAT CCG CAA GAT GGG ATT ATG GTG Asp Ile Met Arg Ala Ser Gln Asp Asn Pro Gln Asp Gly Ile Met Val 1630 1635 1640 1645	5124
AAA CTA GTT GTC AAT TTG TTG CAG TTA TCC AAG ATG GCA ATA AAC CAC Lys Leu Val Val Asn Leu Leu Gln Leu Ser Lys Met Ala Ile Asn His 1650 1655 1660	5172
ACT GGT GAA AAA GAA GTT CTA GAG GCT GTT GGA AGC TGC TTG GGA GAA Thr Gly Glu Lys Glu Val Leu Glu Ala Val Gly Ser Cys Leu Gly Glu 1665 1670 1675	5220
GTG GGT CCT ATA GAT TTC TCT ACC ATA GCT ATA CAA CAT AGT AAA GAT Val Gly Pro Ile Asp Phe Ser Thr Ile Ala Ile Gln His Ser Lys Asp 1680 1685 1690	5268
GCA TCT TAT ACC AAG GCC CTT AAG TTA TTT GAA GAT AAA GAA CTT CAG Ala Ser Tyr Thr Lys Ala Leu Lys Leu Phe Glu Asp Lys Glu Leu Gln 1695 1700 1705	5316
TGG ACC TTC ATA ATG CTG ACC TAC CTG AAT AAC ACA CTG GTA GAA GAT Trp Thr Phe Ile Met Leu Thr Tyr Leu Asn Asn Thr Leu Val Glu Asp 1710 1715 1720 1725	5364
TGT GTC AAA GTT CGA TCA GCA GCT GTT ACC TGT TTG AAA AAC ATT TTA Cys Val Lys Val Arg Ser Ala Ala Val Thr Cys Leu Lys Asn Ile Leu 1730 1735 1740	5412
GCC ACA AAG ACT GGA CAT AGT TTC TGG GAG ATT TAT AAG ATG ACA ACA Ala Thr Lys Thr Gly His Ser Phe Trp Glu Ile Tyr Lys Met Thr Thr 1745 1750 1755	5460
GAT CCA ATG CTG GCC TAT CTA CAG CCT TTT AGA ACA TCA AGA AAA AAG Asp Pro Met Leu Ala Tyr Leu Gln Pro Phe Arg Thr Ser Arg Lys Lys 1760 1765 1770	5508
TTT TTA GAA GTA CCC AGA TTT GAC AAA GAA AAC CCT TTT GAA GGC CTG Phe Leu Glu Val Pro Arg Phe Asp Lys Glu Asn Pro Phe Glu Gly Leu 1775 1780 1785	5556
GAT GAT ATA AAT CTG TGG ATT CCT CTA AGT GAA AAT CAT GAC ATT TGG Asp Asp Ile Asn Leu Trp Ile Pro Leu Ser Glu Asn His Asp Ile Trp 1790 1795 1800 1805	5604
ATA AAG ACA CTG ACT TGT GCT TTT TTG GAC AGT GGA GGC ACA AAA TGT Ile Lys Thr Leu Thr Cys Ala Phe Leu Asp Ser Gly Gly Thr Lys Cys 1810 1815 1820	5652
GAA ATT CTT CAA TTA TTA AAG CCA ATG TGT GAA GTG AAA ACT GAC TTT Glu Ile Leu Gln Leu Leu Lys Pro Met Cys Glu Val Lys Thr Asp Phe 1825 1830 1835	5700
TGT CAG ACT GTA CTT CCA TAC TTG ATT CAT GAT ATT TTA CTC CAA GAT Cys Gln Thr Val Leu Pro Tyr Leu Ile His Asp Ile Leu Leu Gln Asp 1840 1845 1850	5748
ACA AAT GAA TCA TGG AGA AAT CTG CTT TCT ACA CAT GTT CAG GGA TTT Thr Asn Glu Ser Trp Arg Asn Leu Leu Ser Thr His Val Gln Gly Phe 1855 1860 1865	5796
TTC ACC AGC TGT CTT CGA CAC TTC TCG CAA ACG AGC CGA TCC ACA ACC Phe Thr Ser Cys Leu Arg His Phe Ser Gln Thr Ser Arg Ser Thr Thr 1870 1875 1880 1885	5844
CCT GCA AAC TTG GAT TCA GAG TCA GAG CAC TTT TTC CGA TGC TGT TTG Pro Ala Asn Leu Asp Ser Glu Ser Glu His Phe Phe Arg Cys Cys Leu 1890 1895 1900	5892

-189-

GAT AAA AAA TCA CAA AGA ACA ATG CTT GCT GTT GTG GAC TAC ATG AGA Asp Lys Lys Ser Gln Arg Thr Met Leu Ala Val Val Asp Tyr Met Arg 1905 1910 1915	5940
AGA CAA AAG AGA CCT TCT TCA GGA ACA ATT TTT AAT GAT GCT TTC TGG Arg Gln Lys Arg Pro Ser Ser Gly Thr Ile Phe Asn Asp Ala Phe Trp 1920 1925 1930	5988
CTG GAT TTA AAT TAT CTA GAA GTT GCC AAG GTA GCT CAG TCT TGT GCT Leu Asp Leu Asn Tyr Leu Glu Val Ala Lys Val Ala Gln Ser Cys Ala 1935 1940 1945	6036
GCT CAC TTT ACA GCT TTA CTC TAT GCA GAA ATC TAT GCA GAT AAG AAA Ala His Phe Thr Ala Leu Leu Tyr Ala Glu Ile Tyr Ala Asp Lys Lys 1950 1955 1960 1965	6084
AGT ATG GAT GAT CAA GAG AAA AGA AGT CTT GCA TTT GAA GAA GGA AGC Ser Met Asp Asp Gln Glu Lys Arg Ser Leu Ala Phe Glu Glu Gly Ser 1970 1975 1980	6132
CAG AGT ACA ACT ATT TCT AGC TTG AGT GAA AAA AGT AAA GAA GAA ACT Gln Ser Thr Thr Ile Ser Ser Leu Ser Glu Lys Ser Lys Glu Glu Thr 1985 1990 1995	6180
GGA ATA AGT TTA CAG GAT CTT CTC TTA GAA ATC TAC AGA AGT ATA GGG Gly Ile Ser Leu Gln Asp Leu Leu Leu Glu Ile Tyr Arg Ser Ile Gly 2000 2005 2010	6228
GAG CCA GAT AGT TTG TAT GGC TGT GGT GGA GGG AAG ATG TTA CAA CCC Glu Pro Asp Ser Leu Tyr Gly Cys Gly Gly Lys Met Leu Gln Pro 2015 2020 2025	6276
ATT ACT AGA CTA CGA ACA TAT GAA CAC GAA GCA ATG TGG GGC AAA GCC Ile Thr Arg Leu Arg Thr Tyr Glu His Glu Ala Met Trp Gly Lys Ala 2030 2035 2040 2045	6324
CTA GTA ACA TAT GAC CTC GAA ACA GCA ATC CCC TCA TCA ACA CGC CAG Leu Val Thr Tyr Asp Leu Glu Thr Ala Ile Pro Ser Ser Thr Arg Gln 2050 2055 2060	6372
GCA GGA ATC ATT CAG GCC TTG CAG AAT TTG GGA CTC TGC CAT ATT CTT Ala Gly Ile Ile Gln Ala Leu Gln Asn Leu Gly Leu Cys His Ile Leu 2065 2070 2075	6420
TCC GTC TAT TTA AAA GGA TTG GAT TAT GAA AAT AAA GAC TGG TGT CCT Ser Val Tyr Leu Lys Gly Leu Asp Tyr Glu Asn Lys Asp Trp Cys Pro 2080 2085 2090	6468
GAA CTA GAA GAA CTT CAT TAC CAA GCA GCA TGG AGG AAT ATG CAG TGG Glu Leu Glu Glu Leu His Tyr Gln Ala Ala Trp Arg Asn Met Gln Trp 2095 2100 2105	6516
GAC CAT TGC ACT TCC GTC AGC AAA GAA GTA GAA GGA ACC AGT TAC CAT Asp His Cys Thr Ser Val Ser Lys Glu Val Glu Gly Thr Ser Tyr His 2110 2115 2120 2125	6564
GAA TCA TTG TAC AAT GCT CTA CAA TCT CTA AGA GAC AGA GAA TTC TCT Glu Ser Leu Tyr Asn Ala Leu Gln Ser Leu Arg Asp Arg Glu Phe Ser 2130 2135 2140	6612
ACA TTT TAT GAA AGT CTC AAA TAT GCC AGA GTA AAA GAA GTG GAA GAG Thr Phe Tyr Glu Ser Leu Lys Tyr Ala Arg Val Lys Glu Val Glu Glu 2145 2150 2155	6660
ATG TGT AAG CGC AGC CTT GAG TCT GTG TAT TCG CTC TAT CCC ACA CTT Met Cys Lys Arg Ser Leu Glu Ser Val Tyr Ser Leu Tyr Pro Thr Leu 2160 2165 2170	6708

-190-

AGC AGG TTG CAG GCC ATT GGA GAG CTG GAA AGC ATT GGG GAG CTT TTC Ser Arg Leu Gln Ala Ile Gly Glu Leu Glu Ser Ile Gly Glu Leu Phe 2175 2180 2185	6756
TCA AGA TCA GTC ACA CAT AGA CAA CTC TCT GAA GTA TAT ATT AAG TGG Ser Arg Ser Val Thr His Arg Gln Leu Ser Glu Val Tyr Ile Lys Trp 2190 2195 2200 2205	6804
CAG AAA CAC TCC CAG CTT CTC AAG GAC AGT GAT TTT AGT TTT CAG GAG Gln Lys His Ser Gln Leu Leu Lys Asp Ser Asp Phe Ser Phe Gln Glu 2210 2215 2220	6852
CCT ATC ATG GCT CTA CGC ACA GTC ATT TTG GAG ATC CTG ATG GAA AAG Pro Ile Met Ala Leu Arg Thr Val Ile Leu Glu Ile Leu Met Glu Lys 2225 2230 2235	6900
GAA ATG GAC AAC TCA CAA AGA GAA TGT ATT AAG GAC ATT CTC ACC AAA Glu Met Asp Asn Ser Gln Arg Glu Cys Ile Lys Asp Ile Leu Thr Lys 2240 2245 2250	6948
CAC CTT GTA GAA CTC TCT ATA CTG GCC AGA ACT TTC AAG AAC ACT CAG His Leu Val Glu Leu Ser Ile Leu Ala Arg Thr Phe Lys Asn Thr Gln 2255 2260 2265	6996
CTC CCT GAA AGG GCA ATA TTT CAA ATT AAA CAG TAC AAT TCA GTT AGC Leu Pro Glu Arg Ala Ile Phe Gln Ile Lys Gln Tyr Asn Ser Val Ser 2270 2275 2280 2285	7044
TGT GGA GTC TCT GAG TGG CAG CTG GAA GAA GCA CAA GTA TTC TGG GCA Cys Gly Val Ser Glu Trp Gln Leu Glu Glu Ala Gln Val Phe Trp Ala 2290 2295 2300	7092
AAA AAG GAG CAG AGT CTT GCC CTG AGT ATT CTC AAG CAA ATG ATC AAG Lys Lys Glu Gln Ser Leu Ala Leu Ser Ile Leu Lys Gln Met Ile Lys 2305 2310 2315	7140
AAG TTG GAT GCC AGC TGT GCA GCG AAC AAT CCC AGC CTA AAA CTT ACA Lys Leu Asp Ala Ser Cys Ala Ala Asn Asn Pro Ser Leu Lys Leu Thr 2320 2325 2330	7188
TAC ACA GAA TGT CTG AGG GTT TGT GGC AAC TGG TTA GCA GAA ACG TGC Tyr Thr Glu Cys Leu Arg Val Cys Gly Asn Trp Leu Ala Glu Thr Cys 2335 2340 2345	7236
TTA GAA AAT CCT GCG GTC ATC ATG CAG ACC TAT CTA GAA AAG GCA GTA Leu Glu Asn Pro Ala Val Ile Met Gln Thr Tyr Leu Glu Lys Ala Val 2350 2355 2360 2365	7284
GAA GTT GCT GGA AAT TAT GAT GGA GAA AGT AGT GAT GAG CTA AGA AAT Glu Val Ala Gly Asn Tyr Asp Gly Glu Ser Ser Asp Glu Leu Arg Asn 2370 2375 2380	7332
GGA AAA ATG AAG GCA TTT CTC TCA TTA GCC CGG TTT TCA GAT ACT CAA Gly Lys Met Lys Ala Phe Leu Ser Leu Ala Arg Phe Ser Asp Thr Gln 2385 2390 2395	7380
TAC CAA AGA ATT GAA AAC TAC ATG AAA TCA TCG GAA TTT GAA AAC AAG Tyr Gln Arg Ile Glu Asn Tyr Met Lys Ser Ser Glu Phe Glu Asn Lys 2400 2405 2410	7428
CAA GCT CTC CTG AAA AGA GCC AAA GAG GAA GTA GGT CTC CTT AGG GAA Gln Ala Leu Leu Lys Arg Ala Lys Glu Glu Val Gly Leu Leu Arg Glu 2415 2420 2425	7476
CAT AAA ATT CAG ACA AAC AGA TAC ACA GTA AAG GTT CAG CGA GAG CTG His Lys Ile Gln Thr Asn Arg Tyr Thr Val Lys Val Gln Arg Glu Leu 2430 2435 2440 2445	7524

-191-

GAG TTG GAT GAA TTA GCC CTG CGT GCA CTG AAA GAG GAT CGT AAA CGC Glu Leu Asp Glu Leu Ala Leu Arg Ala Leu Lys Glu Asp Arg Lys Arg 2450 2455 2460	7572
TTC TTA TGT AAA GCA GTT GAA AAT TAT ATC AAC TGC TTA TTA AGT GGA Phe Leu Cys Lys Ala Val Glu Asn Tyr Ile Asn Cys Leu Leu Ser Gly 2465 2470 2475	7620
GAA GAA CAT GAT ATG TGG GTA TTC CGG CTT TGT TCC CTC TGG CTT GAA Glu Glu His Asp Met Trp Val Phe Arg Leu Cys Ser Leu Trp Leu Glu 2480 2485 2490	7668
AAT TCT GGA GTT TCT GAA GTC AAT GGC ATG ATG AAG AGA GAC GGA ATG Asn Ser Gly Val Ser Glu Val Asn Gly Met Met Lys Arg Asp Gly Met 2495 2500 2505	7716
AAG ATT CCA ACA TAT AAA TTT TTG CCT CTT ATG TAC CAA TTG GCT GCT Lys Ile Pro Thr Tyr Lys Phe Leu Pro Leu Met Tyr Gln Leu Ala Ala 2510 2515 2520 2525	7764
AGA ATG GGG ACC AAG ATG ATG GGA GGC CTA GGA TTT CAT GAA GTC CTC Arg Met Gly Thr Lys Met Met Gly Gly Leu Gly Phe His Glu Val Leu 2530 2535 2540	7812
AAT AAT CTA ATC TCT AGA ATT TCA ATG GAT CAC CCC CAT CAC ACT TTG Asn Asn Leu Ile Ser Arg Ile Ser Met Asp His Pro His His Thr Leu 2545 2550 2555	7860
TTT ATT ATA CTG GCC TTA GCA AAT GCA AAC AGA GAT GAA TTT CTG ACT Phe Ile Leu Ala Leu Ala Asn Ala Asn Arg Asp Glu Phe Leu Thr 2560 2565 2570	7908
AAA CCA GAG GTA GCC AGA AGA AGC AGA ATA ACT AAA AAT GTG CCT AAA Lys Pro Glu Val Ala Arg Arg Ser Arg Ile Thr Lys Asn Val Pro Lys 2575 2580 2585	7956
CAA AGC TCT CAG CTT GAT GAG GAT CGA ACA GAG GCT GCA AAT AGA ATA Gln Ser Ser Gln Leu Asp Glu Asp Arg Thr Glu Ala Ala Asn Arg Ile 2590 2595 2600 2605	8004
ATA TGT ACT ATC AGA AGT AGG AGA CCT CAG ATG GTC AGA AGT GTT GAG Ile Cys Thr Ile Arg Ser Arg Arg Pro Gln Met Val Arg Ser Val Glu 2610 2615 2620	8052
GCA CTT TGT GAT GCT TAT ATT ATA TTA GCA AAC TTA GAT GCC ACT CAG Ala Leu Cys Asp Ala Tyr Ile Ile Leu Ala Asn Leu Asp Ala Thr Gln 2625 2630 2635	8100
TGG AAG ACT CAG AGA AAA GGC ATA AAT ATT CCA GCA GAC CAG CCA ATT Trp Lys Thr Gln Arg Lys Gly Ile Asn Ile Pro Ala Asp Gln Pro Ile 2640 2645 2650	8148
ACT AAA CTT AAG AAT TTA GAA GAT GTT GTT GTC CCT ACT ATG GAA ATT Thr Lys Leu Lys Asn Leu Glu Asp Val Val Val Pro Thr Met Glu Ile 2655 2660 2665	8196
AAG GTG GAC CAC ACA GGA GAA TAT GGA AAT CTG GTG ACT ATA CAG TCA Lys Val Asp His Thr Gly Glu Tyr Gly Asn Leu Val Thr Ile Gln Ser 2670 2675 2680 2685	8244
TTT AAA GCA GAA TTT CGC TTA GCA GGA GGT GTA AAT TTA CCA AAA ATA Phe Lys Ala Glu Phe Arg Leu Ala Gly Gly Val Asn Leu Pro Lys Ile 2690 2695 2700	8292
ATA GAT TGT GTA GGT TCC GAT GGC AAG GAG AGG AGA CAG CTT GTT AAG Ile Asp Cys Val Gly Ser Asp Gly Lys Glu Arg Arg Gln Leu Val Lys 2705 2710 2715	8340

-192-

GGC CGT GAT GAC CTG AGA CAA GAT GCT GTC ATG CAA CAG GTC TTC CAG	8388
Gly Arg Asp Asp Leu Arg Gln Asp Ala Val Met Gln Gln Val Phe Gln	
2720 2725 2730	
ATG TGT AAT ACA TTA CTG CAG AGA AAC ACG GAA ACT AGG AAG AGG AAA	8436
Met Cys Asn Thr Leu Leu Gln Arg Asn Thr Glu Thr Arg Lys Arg Lys	
2735 2740 2745	
TTA ACT ATC TGT ACT TAT AAG GTG GTT CCC CTC TCT CAG CGA AGT GGT	8484
Leu Thr Ile Cys Thr Tyr Lys Val Val Pro Leu Ser Gln Arg Ser Gly	
2750 2755 2760 2765	
GTT CTT GAA TGG TGC ACA GGA ACT GTC CCC ATT GGT GAA TTT CTT GTT	8532
Val Leu Glu Trp Cys Thr Gly Thr Val Pro Ile Gly Glu Phe Leu Val	
2770 2775 2780	
AAC AAT GAA GAT GGT GCT CAT AAA AGA TAC AGG CCA AAT GAT TTC AGT	8580
Asn Asn Glu Asp Gly Ala His Lys Arg Tyr Arg Pro Asn Asp Phe Ser	
2785 2790 2795	
GCC TTT CAG TGC CAA AAG AAA ATG ATG GAG GTG CAA AAA AAG TCT TTT	8628
Ala Phe Gln Cys Gln Lys Lys Met Met Glu Val Gln Lys Lys Ser Phe	
2800 2805 2810	
GAA GAG AAA TAT GAA GTC TTC ATG GAT GTT TGC CAA AAT TTT CAA CCA	8676
Glu Glu Lys Tyr Glu Val Phe Met Asp Val Cys Gln Asn Phe Gln Pro	
2815 2820 2825	
GTT TTC CGT TAC TTC TGC ATG GAA AAA TTC TTG GAT CCA GCT ATT TGG	8724
Val Phe Arg Tyr Phe Cys Met Glu Lys Phe Leu Asp Pro Ala Ile Trp	
2830 2835 2840 2845	
TTT GAG AAG CGA TTG GCT TAT ACG CGC AGT GTA GCT ACT TCT TCT ATT	8772
Phe Glu Lys Arg Leu Ala Tyr Thr Arg Ser Val Ala Thr Ser Ser Ile	
2850 2855 2860	
GTT GGT TAC ATA CTT GGA CTT GGT GAT AGA CAT GTA CAG AAT ATC TTG	8820
Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Val Gln Asn Ile Leu	
2865 2870 2875	
ATA AAT GAG CAG TCA GCA GAA CTT GTA CAT ATA GAT CTA GGT GTT GCT	8868
Ile Asn Glu Gln Ser Ala Glu Leu Val His Ile Asp Leu Gly Val Ala	
2880 2885 2890	
TTT GAA CAG GGC AAA ATC CTT CCT ACT CCT GAG ACA GTT CCT TTT AGA	8916
Phe Glu Gln Gly Lys Ile Leu Pro Thr Pro Glu Thr Val Pro Phe Arg	
2895 2900 2905	
CTC ACC AGA GAT ATT GTG GAT GGC ATG GGC ATT ACG GGT GTT GAA GGT	8964
Leu Thr Arg Asp Ile Val Asp Gly Met Gly Ile Thr Gly Val Glu Gly	
2910 2915 2920 2925	
GTC TTC AGA AGA TGC TGT GAG AAA ACC ATG GAA GTG ATG AGA AAC TCT	9012
Val Phe Arg Arg Cys Cys Glu Lys Thr Met Glu Val Met Arg Asn Ser	
2930 2935 2940	
CAG GAA ACT CTG TTA ACC ATT GTA GAG GTC CTT CTA TAT GAT CCA CTC	9060
Gln Glu Thr Leu Leu Thr Ile Val Glu Val Leu Leu Tyr Asp Pro Leu	
2945 2950 2955	
TTT GAC TGG ACC ATG AAT CCT TTG AAA GCT TTG TAT TTA CAG CAG AGG	9108
Phe Asp Trp Thr Met Asn Pro Leu Lys Ala Leu Tyr Leu Gln Gln Arg	
2960 2965 2970	
CCG GAA GAT GAA ACT GAG CTT CAC CCT ACT CTG AAT GCA GAT GAC CAA	9156
Pro Glu Asp Glu Thr Glu Leu His Pro Thr Leu Asn Ala Asp Asp Gln	
2975 2980 2985	

-193-

GAA TGC AAA CGA AAT CTC AGT GAT ATT GAC CAG AGT TTC GAC AAA GTA Glu Cys Lys Arg Asn Leu Ser Asp Ile Asp Gln Ser Phe Asp Lys Val 2990 2995 3000 3005	9204
GCT GAA CGT GTC TTA ATG AGA CTA CAA GAG AAA CTG AAA GGA GTG GAA Ala Glu Arg Val Leu Met Arg Leu Gln Glu Lys Leu Lys Gly Val Glu 3010 3015 3020	9252
GAA GGC ACT GTG CTC AGT GTT GGT GGA CAG GTG AAT TTG CTC ATA CAG Glu Gly Thr Val Leu Ser Val Gly Gly Gln Val Asn Leu Leu Ile Gln 3025 3030 3035	9300
CAG GCC ATA GAC CCC AAA AAT CTC AGC CGA CTT TTC CCA GGA TGG AAA Gln Ala Ile Asp Pro Lys Asn Leu Ser Arg Leu Phe Pro Gly Trp Lys 3040 3045 3050	9348
GCT TGG GTG TGATCTTCAG TATATGAATT ACCCTTTC Ala Trp Val 3055	9385

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3056 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Ser Leu Val Leu Asn Asp Leu Leu Ile Cys Cys Arg Gln Leu Glu 1 5 10 15
His Asp Arg Ala Thr Glu Arg Lys Lys Glu Val Glu Lys Phe Lys Arg 20 25 30
Leu Ile Arg Asp Pro Glu Thr Ile Lys His Leu Asp Arg His Ser Asp 35 40 45
Ser Lys Lys Gly Lys Tyr Leu Asn Trp Asp Ala Val Phe Arg Phe Leu 50 55 60
Gln Lys Tyr Ile Gln Lys Glu Thr Glu Cys Leu Arg Ile Ala Lys Pro 65 70 75 80
Asn Val Ser Ala Ser Thr Gln Ala Ser Arg Gln Lys Lys Met Gln Glu 85 90 95
Ile Ser Ser Leu Val Lys Tyr Phe Ile Lys Cys Ala Asn Arg Arg Ala 100 105 110
Pro Arg Leu Lys Cys Gln Glu Leu Leu Asn Tyr Ile Met Asp Thr Val 115 120 125
Lys Asp Ser Ser Asn Gly Ala Ile Tyr Gly Ala Asp Cys Ser Asn Ile 130 135 140
Leu Leu Lys Asp Ile Leu Ser Val Arg Lys Tyr Trp Cys Glu Ile Ser 145 150 155 160
Gln Gln Gln Trp Leu Glu Leu Phe Ser Val Tyr Phe Arg Leu Tyr Leu 165 170 175
Lys Pro Ser Gln Asp Val His Arg Val Leu Val Ala Arg Ile Ile His 180 185 190

-194-

Ala Val Thr Lys Gly Cys Cys Ser Gln Thr Asp Gly Leu Asn Ser Lys
 195 200 205
 Phe Leu Asp Phe Phe Ser Lys Ala Ile Gln Cys Ala Arg Gln Glu Lys
 210 215 220
 Ser Ser Ser Gly Leu Asn His Ile Leu Ala Ala Leu Thr Ile Phe Leu
 225 230 235 240
 Lys Thr Leu Ala Val Asn Phe Arg Ile Arg Val Cys Glu Leu Gly Asp
 245 250 255
 Glu Ile Leu Pro Thr Leu Leu Tyr Ile Trp Thr Gln His Arg Leu Asn
 260 265 270
 Asp Ser Leu Lys Glu Val Ile Ile Glu Leu Phe Gln Leu Gln Ile Tyr
 275 280 285
 Ile His His Pro Lys Gly Ala Lys Thr Gln Glu Lys Gly Ala Tyr Glu
 290 295 300
 Ser Thr Lys Trp Arg Ser Ile Leu Tyr Asn Leu Tyr Asp Leu Leu Val
 305 310 315 320
 Asn Glu Ile Ser His Ile Gly Ser Arg Gly Lys Tyr Ser Ser Gly Phe
 325 330 335
 Arg Asn Ile Ala Val Lys Glu Asn Leu Ile Glu Leu Met Ala Asp Ile
 340 345 350
 Cys His Gln Val Phe Asn Glu Asp Thr Arg Ser Leu Glu Ile Ser Gln
 355 360 365
 Ser Tyr Thr Thr Thr Gln Arg Glu Ser Ser Asp Tyr Ser Val Pro Cys
 370 375 380
 Lys Arg Lys Lys Ile Glu Leu Gly Trp Glu Val Ile Lys Asp His Leu
 385 390 395 400
 Gln Lys Ser Gln Asn Asp Phe Asp Leu Val Pro Trp Leu Gln Ile Ala
 405 410 415
 Thr Gln Leu Ile Ser Lys Tyr Pro Ala Ser Leu Pro Asn Cys Glu Leu
 420 425 430
 Ser Pro Leu Leu Met Ile Leu Ser Gln Leu Leu Pro Gln Gln Arg His
 435 440 445
 Gly Glu Arg Thr Pro Tyr Val Leu Arg Cys Leu Thr Glu Val Ala Leu
 450 455 460
 Cys Gln Asp Lys Arg Ser Asn Leu Glu Ser Ser Gln Lys Ser Asp Leu
 465 470 475 480
 Leu Lys Leu Trp Asn Lys Ile Trp Cys Ile Thr Phe Arg Gly Ile Ser
 485 490 495
 Ser Glu Gln Ile Gln Ala Glu Asn Phe Gly Leu Leu Gly Ala Ile Ile
 500 505 510
 Gln Gly Ser Leu Val Glu Val Asp Arg Glu Phe Trp Lys Leu Phe Thr
 515 520 525
 Gly Ser Ala Cys Arg Pro Ser Cys Pro Ala Val Cys Cys Leu Thr Leu
 530 535 540

-195-

Ala Leu Thr Thr Ser Ile Val Pro Gly Ala Val Lys Met Gly Ile Glu
 545 550 555 560
 Gln Asn Met Cys Glu Val Asn Arg Ser Phe Ser Leu Lys Glu Ser Ile
 565 570 575
 Met Lys Trp Leu Leu Phe Tyr Gln Leu Glu Gly Asp Leu Glu Asn Ser
 580 585 590
 Thr Glu Val Pro Pro Ile Leu His Ser Asn Phe Pro His Leu Val Leu
 595 600 605
 Glu Lys Ile Leu Val Ser Leu Thr Met Lys Asn Cys Lys Ala Ala Met
 610 615 620
 Asn Phe Phe Gln Ser Val Pro Glu Cys Glu His His Gln Lys Asp Lys
 625 630 635 640
 Glu Glu Leu Ser Phe Ser Glu Val Glu Glu Leu Phe Leu Gln Thr Thr
 645 650 655
 Phe Asp Lys Met Asp Phe Leu Thr Ile Val Arg Glu Cys Gly Ile Glu
 660 665 670
 Lys His Gln Ser Ser Ile Gly Phe Ser Val His Gln Asn Leu Lys Glu
 675 680 685
 Ser Leu Asp Arg Cys Leu Leu Gly Leu Ser Glu Gln Leu Leu Asn Asn
 690 695 700
 Tyr Ser Ser Glu Ile Thr Asn Ser Glu Thr Leu Val Arg Cys Ser Arg
 705 710 715 720
 Leu Leu Val Gly Val Leu Gly Cys Tyr Cys Tyr Met Gly Val Ile Ala
 725 730 735
 Glu Glu Glu Ala Tyr Lys Ser Glu Leu Phe Gln Lys Ala Asn Ser Leu
 740 745 750
 Met Gln Cys Ala Gly Glu Ser Ile Thr Leu Phe Lys Asn Lys Thr Asn
 755 760 765
 Glu Glu Phe Arg Ile Gly Ser Leu Arg Asn Met Met Gln Leu Cys Thr
 770 775 780
 Arg Cys Leu Ser Asn Cys Thr Lys Lys Ser Pro Asn Lys Ile Ala Ser
 785 790 795 800
 Gly Phe Phe Leu Arg Leu Leu Thr Ser Lys Leu Met Asn Asp Ile Ala
 805 810 815
 Asp Ile Cys Lys Ser Leu Ala Ser Phe Ile Lys Lys Pro Phe Asp Arg
 820 825 830
 Gly Glu Val Glu Ser Met Glu Asp Asp Thr Asn Gly Asn Leu Met Glu
 835 840 845
 Val Glu Asp Gln Ser Ser Met Asn Leu Phe Asn Asp Tyr Pro Asp Ser
 850 855 860
 Ser Val Ser Asp Ala Asn Glu Pro Gly Glu Ser Gln Ser Thr Ile Gly
 865 870 875 880
 Ala Ile Asn Pro Leu Ala Glu Glu Tyr Leu Ser Lys Gln Asp Leu Leu
 885 890 895

-196-

Phe	Leu	Asp	Met	Leu	Lys	Phe	Leu	Cys	Leu	Cys	Val	Thr	Thr	Ala	Gln
			900					905						910	
Thr	Asn	Thr	Val	Ser	Phe	Arg	Ala	Ala	Asp	Ile	Arg	Arg	Lys	Leu	Leu
		915					920					925			
Met	Leu	Ile	Asp	Ser	Ser	Thr	Leu	Glu	Pro	Thr	Lys	Ser	Leu	His	Leu
	930					935					940				
His	Met	Tyr	Leu	Met	Leu	Leu	Lys	Glu	Leu	Pro	Gly	Glu	Glu	Tyr	Pro
945					950					955					960
Leu	Pro	Met	Glu	Asp	Val	Leu	Glu	Leu	Leu	Lys	Pro	Leu	Ser	Asn	Val
				965					970					975	
Cys	Ser	Leu	Tyr	Arg	Arg	Asp	Gln	Asp	Val	Cys	Lys	Thr	Ile	Leu	Asn
			980					985					990		
His	Val	Leu	His	Val	Val	Lys	Asn	Leu	Gly	Gln	Ser	Asn	Met	Asp	Ser
		995					1000					1005			
Glu	Asn	Thr	Arg	Asp	Ala	Gln	Gly	Gln	Phe	Leu	Thr	Val	Ile	Gly	Ala
	1010					1015						1020			
Phe	Trp	His	Leu	Thr	Lys	Glu	Arg	Lys	Tyr	Ile	Phe	Ser	Val	Arg	Met
1025					1030					1035					1040
Ala	Leu	Val	Asn	Cys	Leu	Lys	Thr	Leu	Leu	Glu	Ala	Asp	Pro	Tyr	Ser
				1045					1050					1055	
Lys	Trp	Ala	Ile	Leu	Asn	Val	Met	Gly	Lys	Asp	Phe	Pro	Val	Asn	Glu
			1060					1065					1070		
Val	Phe	Thr	Gln	Phe	Leu	Ala	Asp	Asn	His	His	Gln	Val	Arg	Met	Leu
		1075					1080					1085			
Ala	Ala	Glu	Ser	Ile	Asn	Arg	Leu	Phe	Gln	Asp	Thr	Lys	Gly	Asp	Ser
	1090					1095					1100				
Ser	Arg	Leu	Leu	Lys	Ala	Leu	Pro	Leu	Lys	Leu	Gln	Gln	Thr	Ala	Phe
1105					1110					1115					1120
Glu	Asn	Ala	Tyr	Leu	Lys	Ala	Gln	Glu	Gly	Met	Arg	Glu	Met	Ser	His
				1125					1130					1135	
Ser	Ala	Glu	Asn	Pro	Glu	Thr	Leu	Asp	Glu	Ile	Tyr	Asn	Arg	Lys	Ser
			1140					1145					1150		
Val	Leu	Leu	Thr	Leu	Ile	Ala	Val	Val	Leu	Ser	Cys	Ser	Pro	Ile	Cys
		1155					1160					1165			
Glu	Lys	Gln	Ala	Leu	Phe	Ala	Leu	Cys	Lys	Ser	Val	Lys	Glu	Asn	Gly
	1170					1175					1180				
Leu	Glu	Pro	His	Leu	Val	Lys	Lys	Val	Leu	Glu	Lys	Val	Ser	Glu	Thr
1185					1190					1195					1200
Phe	Gly	Tyr	Arg	Arg	Leu	Glu	Asp	Phe	Met	Ala	Ser	His	Leu	Asp	Tyr
			1205					1210						1215	
Leu	Val	Leu	Glu	Trp	Leu	Asn	Leu	Gln	Asp	Thr	Glu	Tyr	Asn	Leu	Ser
			1220					1225							

-197-

Arg Ser Cys Tyr Lys Val Leu Ile Pro His Leu Val Ile Arg Ser His
 1250 1255 1260
 Phe Asp Glu Val Lys Ser Ile Ala Asn Gln Ile Gln Glu Asp Trp Lys
 1265 1270 1275 1280
 Ser Leu Leu Thr Asp Cys Phe Pro Lys Ile Leu Val Asn Ile Leu Pro
 1285 1290 1295
 Tyr Phe Ala Tyr Glu Gly Thr Arg Asp Ser Gly Met Ala Gln Gln Arg
 1300 1305 1310
 Glu Thr Ala Thr Lys Val Tyr Asp Met Leu Lys Ser Glu Asn Leu Leu
 1315 1320 1325
 Gly Lys Gln Ile Asp His Leu Phe Ile Ser Asn Leu Pro Glu Ile Val
 1330 1335 1340
 Val Glu Leu Leu Met Thr Leu His Glu Pro Ala Asn Ser Ser Ala Ser
 1345 1350 1355 1360
 Gln Ser Thr Asp Leu Cys Asp Phe Ser Gly Asp Leu Asp Pro Ala Pro
 1365 1370 1375
 Asn Pro Pro His Phe Pro Ser His Val Ile Lys Ala Thr Phe Ala Tyr
 1380 1385 1390
 Ile Ser Asn Cys His Lys Thr Lys Leu Lys Ser Ile Leu Glu Ile Leu
 1395 1400 1405
 Ser Lys Ser Pro Asp Ser Tyr Gln Lys Ile Leu Leu Ala Ile Cys Glu
 1410 1415 1420
 Gln Ala Ala Glu Thr Asn Asn Val Tyr Lys Lys His Arg Ile Leu Lys
 1425 1430 1435 1440
 Ile Tyr His Leu Phe Val Ser Leu Leu Leu Lys Asp Ile Lys Ser Gly
 1445 1450 1455
 Leu Gly Gly Ala Trp Ala Phe Val Leu Arg Asp Val Ile Tyr Thr Leu
 1460 1465 1470
 Ile His Tyr Ile Asn Gln Arg Pro Ser Cys Ile Met Asp Val Ser Leu
 1475 1480 1485
 Arg Ser Phe Ser Leu Cys Cys Asp Leu Leu Ser Gln Val Cys Gln Thr
 1490 1495 1500
 Ala Val Thr Tyr Cys Lys Asp Ala Leu Glu Asn His Leu His Val Ile
 1505 1510 1515 1520
 Val Gly Thr Leu Ile Pro Leu Val Tyr Glu Gln Val Glu Val Gln Lys
 1525 1530 1535
 Gln Val Leu Asp Leu Leu Lys Tyr Leu Val Ile Asp Asn Lys Asp Asn
 1540 1545 1550
 Glu Asn Leu Tyr Ile Thr Ile Lys Leu Leu Asp Pro Phe Pro Asp His
 1555 1560 1565
 Val Val Phe Lys Asp Leu Arg Ile Thr Gln Gln Lys Ile Lys Tyr Ser
 1570 1575 1580
 Arg Gly Pro Phe Ser Leu Leu Glu Glu Ile Asn His Phe Leu Ser Val
 1585 1590 1595 1600

-198-

Ser Val Tyr Asp Ala Leu Pro Leu Thr Arg Leu Glu Gly Leu Lys Asp
 1605 1610 1615
 Leu Arg Arg Gln Leu Glu Leu His Lys Asp Gln Met Val Asp Ile Met
 1620 1625 1630
 Arg Ala Ser Gln Asp Asn Pro Gln Asp Gly Ile Met Val Lys Leu Val
 1635 1640 1645
 Val Asn Leu Leu Gln Leu Ser Lys Met Ala Ile Asn His Thr Gly Glu
 1650 1655 1660
 Lys Glu Val Leu Glu Ala Val Gly Ser Cys Leu Gly Glu Val Gly Pro
 1665 1670 1675 1680
 Ile Asp Phe Ser Thr Ile Ala Ile Gln His Ser Lys Asp Ala Ser Tyr
 1685 1690 1695
 Thr Lys Ala Leu Lys Leu Phe Glu Asp Lys Glu Leu Gln Trp Thr Phe
 1700 1705 1710
 Ile Met Leu Thr Tyr Leu Asn Asn Thr Leu Val Glu Asp Cys Val Lys
 1715 1720 1725
 Val Arg Ser Ala Ala Val Thr Cys Leu Lys Asn Ile Leu Ala Thr Lys
 1730 1735 1740
 Thr Gly His Ser Phe Trp Glu Ile Tyr Lys Met Thr Thr Asp Pro Met
 1745 1750 1755 1760
 Leu Ala Tyr Leu Gln Pro Phe Arg Thr Ser Arg Lys Lys Phe Leu Glu
 1765 1770 1775
 Val Pro Arg Phe Asp Lys Glu Asn Pro Phe Glu Gly Leu Asp Asp Ile
 1780 1785 1790
 Asn Leu Trp Ile Pro Leu Ser Glu Asn His Asp Ile Trp Ile Lys Thr
 1795 1800 1805
 Leu Thr Cys Ala Phe Leu Asp Ser Gly Gly Thr Lys Cys Glu Ile Leu
 1810 1815 1820
 Gln Leu Leu Lys Pro Met Cys Glu Val Lys Thr Asp Phe Cys Gln Thr
 1825 1830 1835 1840
 Val Leu Pro Tyr Leu Ile His Asp Ile Leu Leu Gln Asp Thr Asn Glu
 1845 1850 1855
 Ser Trp Arg Asn Leu Leu Ser Thr His Val Gln Gly Phe Phe Thr Ser
 1860 1865 1870
 Cys Leu Arg His Phe Ser Gln Thr Ser Arg Ser Thr Thr Pro Ala Asn
 1875 1880 1885
 Leu Asp Ser Glu Ser Glu His Phe Phe Arg Cys Cys Leu Asp Lys Lys
 1890 1895 1900
 Ser Gln Arg Thr Met Leu Ala Val Val Asp Tyr Met Arg Arg Gln Lys
 1905 1910 1915 1920
 Arg Pro Ser Ser Gly Thr Ile Phe Asn Asp Ala Phe Trp Leu Asp Leu
 1925 1930 1935
 Asn Tyr Leu Glu Val Ala Lys Val Ala Gln Ser Cys Ala Ala His Phe
 1940 1945 1950

-199-

Thr Ala Leu Leu Tyr Ala Glu Ile Tyr Ala Asp Lys Lys Ser Met Asp
 1955 1960 1965
 Asp Gln Glu Lys Arg Ser Leu Ala Phe Glu Glu Gly Ser Gln Ser Thr
 1970 1975 1980
 Thr Ile Ser Ser Leu Ser Glu Lys Ser Lys Glu Glu Thr Gly Ile Ser
 1985 1990 1995 2000
 Leu Gln Asp Leu Leu Leu Glu Ile Tyr Arg Ser Ile Gly Glu Pro Asp
 2005 2010 2015
 Ser Leu Tyr Gly Cys Gly Gly Gly Lys Met Leu Gln Pro Ile Thr Arg
 2020 2025 2030
 Leu Arg Thr Tyr Glu His Glu Ala Met Trp Gly Lys Ala Leu Val Thr
 2035 2040 2045
 Tyr Asp Leu Glu Thr Ala Ile Pro Ser Ser Thr Arg Gln Ala Gly Ile
 2050 2055 2060
 Ile Gln Ala Leu Gln Asn Leu Gly Leu Cys His Ile Leu Ser Val Tyr
 2065 2070 2075 2080
 Leu Lys Gly Leu Asp Tyr Glu Asn Lys Asp Trp Cys Pro Glu Leu Glu
 2085 2090 2095
 Glu Leu His Tyr Gln Ala Ala Trp Arg Asn Met Gln Trp Asp His Cys
 2100 2105 2110
 Thr Ser Val Ser Lys Glu Val Glu Gly Thr Ser Tyr His Glu Ser Leu
 2115 2120 2125
 Tyr Asn Ala Leu Gln Ser Leu Arg Asp Arg Glu Phe Ser Thr Phe Tyr
 2130 2135 2140
 Glu Ser Leu Lys Tyr Ala Arg Val Lys Glu Val Glu Glu Met Cys Lys
 2145 2150 2155 2160
 Arg Ser Leu Glu Ser Val Tyr Ser Leu Tyr Pro Thr Leu Ser Arg Leu
 2165 2170 2175
 Gln Ala Ile Gly Glu Leu Glu Ser Ile Gly Glu Leu Phe Ser Arg Ser
 2180 2185 2190
 Val Thr His Arg Gln Leu Ser Glu Val Tyr Ile Lys Trp Gln Lys His
 2195 2200 2205
 Ser Gln Leu Leu Lys Asp Ser Asp Phe Ser Phe Gln Glu Pro Ile Met
 2210 2215 2220
 Ala Leu Arg Thr Val Ile Leu Glu Ile Leu Met Glu Lys Glu Met Asp
 2225 2230 2235 2240
 Asn Ser Gln Arg Glu Cys Ile Lys Asp Ile Leu Thr Lys His Leu Val
 2245 2250 2255
 Glu Leu Ser Ile Leu Ala Arg Thr Phe Lys Asn Thr Gln Leu Pro Glu
 2260 2265 2270
 Arg Ala Ile Phe Gln Ile Lys Gln Tyr Asn Ser Val Ser Cys Gly Val
 2275 2280 2285
 Ser Glu Trp Gln Leu Glu Glu Ala Gln Val Phe Trp Ala Lys Lys Glu
 2290 2295 2300

-200-

Gln Ser Leu Ala Leu Ser Ile Leu Lys Gln Met Ile Lys Lys Leu Asp
 2305 2310 2315 2320
 Ala Ser Cys Ala Ala Asn Asn Pro Ser Leu Lys Leu Thr Tyr Thr Glu
 2325 2330 2335
 Cys Leu Arg Val Cys Gly Asn Trp Leu Ala Glu Thr Cys Leu Glu Asn
 2340 2345 2350
 Pro Ala Val Ile Met Gln Thr Tyr Leu Glu Lys Ala Val Glu Val Ala
 2355 2360 2365
 Gly Asn Tyr Asp Gly Glu Ser Ser Asp Glu Leu Arg Asn Gly Lys Met
 2370 2375 2380
 Lys Ala Phe Leu Ser Leu Ala Arg Phe Ser Asp Thr Gln Tyr Gln Arg
 2385 2390 2395 2400
 Ile Glu Asn Tyr Met Lys Ser Ser Glu Phe Glu Asn Lys Gln Ala Leu
 2405 2410 2415
 Leu Lys Arg Ala Lys Glu Glu Val Gly Leu Leu Arg Glu His Lys Ile
 2420 2425 2430
 Gln Thr Asn Arg Tyr Thr Val Lys Val Gln Arg Glu Leu Glu Leu Asp
 2435 2440 2445
 Glu Leu Ala Leu Arg Ala Leu Lys Glu Asp Arg Lys Arg Phe Leu Cys
 2450 2455 2460
 Lys Ala Val Glu Asn Tyr Ile Asn Cys Leu Leu Ser Gly Glu Glu His
 2465 2470 2475 2480
 Asp Met Trp Val Phe Arg Leu Cys Ser Leu Trp Leu Glu Asn Ser Gly
 2485 2490 2495
 Val Ser Glu Val Asn Gly Met Met Lys Arg Asp Gly Met Lys Ile Pro
 2500 2505 2510
 Thr Tyr Lys Phe Leu Pro Leu Met Tyr Gln Leu Ala Ala Arg Met Gly
 2515 2520 2525
 Thr Lys Met Met Gly Gly Leu Gly Phe His Glu Val Leu Asn Asn Leu
 2530 2535 2540
 Ile Ser Arg Ile Ser Met Asp His Pro His His Thr Leu Phe Ile Ile
 2545 2550 2555 2560
 Leu Ala Leu Ala Asn Ala Asn Arg Asp Glu Phe Leu Thr Lys Pro Glu
 2565 2570 2575
 Val Ala Arg Arg Ser Arg Ile Thr Lys Asn Val Pro Lys Gln Ser Ser
 2580 2585 2590
 Gln Leu Asp Glu Asp Arg Thr Glu Ala Ala Asn Arg Ile Ile Cys Thr
 2595 2600 2605
 Ile Arg Ser Arg Arg Pro Gln Met Val Arg Ser Val Glu Ala Leu Cys
 2610 2615 2620
 Asp Ala Tyr Ile Ile Leu Ala Asn Leu Asp Ala Thr Gln Trp Lys Thr
 2625 2630 2635 2640
 Gln Arg Lys Gly Ile Asn Ile Pro Ala Asp Gln Pro Ile Thr Lys Leu
 2645 2650 2655

-201-

Lys Asn Leu Glu Asp Val Val Val Pro Thr Met Glu Ile Lys Val Asp
 2660 2665 2670
 His Thr Gly Glu Tyr Gly Asn Leu Val Thr Ile Gln Ser Phe Lys Ala
 2675 2680 2685
 Glu Phe Arg Leu Ala Gly Gly Val Asn Leu Pro Lys Ile Ile Asp Cys
 2690 2695 2700
 Val Gly Ser Asp Gly Lys Glu Arg Arg Gln Leu Val Lys Gly Arg Asp
 2705 2710 2715 2720
 Asp Leu Arg Gln Asp Ala Val Met Gln Gln Val Phe Gln Met Cys Asn
 2725 2730 2735
 Thr Leu Leu Gln Arg Asn Thr Glu Thr Arg Lys Arg Lys Leu Thr Ile
 2740 2745 2750
 Cys Thr Tyr Lys Val Val Pro Leu Ser Gln Arg Ser Gly Val Leu Glu
 2755 2760 2765
 Trp Cys Thr Gly Thr Val Pro Ile Gly Glu Phe Leu Val Asn Asn Glu
 2770 2775 2780
 Asp Gly Ala His Lys Arg Tyr Arg Pro Asn Asp Phe Ser Ala Phe Gln
 2785 2790 2795 2800
 Cys Gln Lys Lys Met Met Glu Val Gln Lys Lys Ser Phe Glu Glu Lys
 2805 2810 2815
 Tyr Glu Val Phe Met Asp Val Cys Gln Asn Phe Gln Pro Val Phe Arg
 2820 2825 2830
 Tyr Phe Cys Met Glu Lys Phe Leu Asp Pro Ala Ile Trp Phe Glu Lys
 2835 2840 2845
 Arg Leu Ala Tyr Thr Arg Ser Val Ala Thr Ser Ser Ile Val Gly Tyr
 2850 2855 2860
 Ile Leu Gly Leu Gly Asp Arg His Val Gln Asn Ile Leu Ile Asn Glu
 2865 2870 2875 2880
 Gln Ser Ala Glu Leu Val His Ile Asp Leu Gly Val Ala Phe Glu Gln
 2885 2890 2895
 Gly Lys Ile Leu Pro Thr Pro Glu Thr Val Pro Phe Arg Leu Thr Arg
 2900 2905 2910
 Asp Ile Val Asp Gly Met Gly Ile Thr Gly Val Glu Gly Val Phe Arg
 2915 2920 2925
 Arg Cys Cys Glu Lys Thr Met Glu Val Met Arg Asn Ser Gln Glu Thr
 2930 2935 2940
 Leu Leu Thr Ile Val Glu Val Leu Leu Tyr Asp Pro Leu Phe Asp Trp
 2945 2950 2955 2960
 Thr Met Asn Pro Leu Lys Ala Leu Tyr Leu Gln Gln Arg Pro Glu Asp
 2965 2970 2975
 Glu Thr Glu Leu His Pro Thr Leu Asn Ala Asp Asp Gln Glu Cys Lys
 2980 2985 2990
 Arg Asn Leu Ser Asp Ile Asp Gln Ser Phe Asp Lys Val Ala Glu Arg
 2995 3000 3005

-202-

Val Leu Met Arg Leu Gln Glu Lys Leu Lys Gly Val Glu Glu Gly Thr
 3010 3015 3020

Val Leu Ser Val Gly Gly Gln Val Asn Leu Leu Ile Gln Gln Ala Ile
 3025 3030 3035 3040

Asp Pro Lys Asn Leu Ser Arg Leu Phe Pro Gly Trp Lys Ala Trp Val
 3045 3050 3055

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met Ser Gly Gly Ser Ser Cys Gln Thr Pro Ser Arg Ala Ile Pro Ala
 1 5 10 15

Thr Arg Arg

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gly Asp Tyr Ser Thr Thr Pro Gly Gly Thr Leu Phe Ser Thr Thr Pro
 1 5 10 15

Gly Gly Thr Arg Arg
 20

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Glu Cys Arg Asn Ser Pro Val Thr Lys Thr Arg Arg
 1 5 10

-203-

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Gly	Val	Thr	Ser	Pro	Ser	Ser	Asp	Glu	Pro	Arg	Arg
1				5					10		

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met	Glu	Ala	Ser	Gln	Ser	His	Leu	Arg	Arg
1			5					10	

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Arg	Arg	Asn	Ser	Pro	Glu	Asp	Lys	Arg	Ala	Gly	Gly
1				5					10		

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

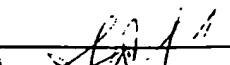
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Gly	Glu	Glu	Ser	Gln	Phe	Glu	Met	Asp	Ile	Arg	Arg
1				5					10		

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>10</u> , lines <u>2-6</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, MD 20852 US	
Date of deposit 7 November 1996	Accession Number s HB 12233 and HB 12234
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>"In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC)."</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
EP	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer  International Division (783) 305-3880	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
---	--

CLAIMS

We claim:

1. A purified and isolated polynucleotide comprising a polynucleotide encoding the PIK-related kinase MCCS1 α amino acid sequence set out in SEQ ID NO: 31.

2. A purified and isolated polynucleotide comprising a polynucleotide encoding the PIK-related kinase MCCS1 β amino acid sequence set out in SEQ ID NO: 33.

3. The polynucleotide of claim 1 or 2 which is a DNA.

4. The DNA of claim 3 which is a cDNA.

5. A MCCS1 α cDNA consisting of the DNA sequence set out in SEQ ID NO: 30.

6. A MCCS1 β DNA consisting of the DNA sequence set out in SEQ ID NO: 32.

7. The DNA of claim 3 which is a genomic DNA.

8. An RNA transcript of the DNA of claim 3.

9. The DNA of claim 3 which is a wholly or partially chemically synthesized DNA.

-206-

10. A DNA comprising a DNA encoding a full length mammalian MCCS1 kinase selected from the group consisting of:

- a) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 30;
- b) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 3; and
- c) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 32.

11. A vector comprising a DNA according to claim 3 or 10.

12. The vector of claim 11 wherein said DNA is operatively linked to an expression control DNA sequence.

13. A host cell stably transformed or transfected with a DNA according to claim 3 or 10 in a manner allowing the expression in said host cell of the MCCS1 kinase.

14. A method for producing the PIK-related kinase MCCS1, said method comprising growing a host cell according to claim 11 in a suitable nutrient medium and isolating the MCCS1 kinase from said cell or the medium of its growth.

15. A purified and isolated polypeptide comprising the PIK-related kinase MCCS1 α amino acid sequence consisting of SEQ ID NO: 31.

16. A purified and isolated polypeptide comprising the PIK-related kinase MCCS1 β amino acid sequence consisting of SEQ ID NO: 33.

17. A polypeptide or peptide capable of specifically binding to PIK-related kinase MCCS1.

18. An antibody product according to claim 17.

-207-

19. A monoclonal antibody according to claim 18.
20. A hybridoma cell line producing a monoclonal antibody according to claim 19.
21. An assay for identifying modulators of MCCS1 kinase activity comprising the steps of:
 - a) incubating a MCCS1 kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate in the presence and absence of a test compound, and
 - b) measuring the moles of phosphate transferred to said substrate; wherein an increase in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an activator of said MCCS1 kinase and a decrease in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an inhibitor of said MCCS1 kinase.
22. The hybridoma cell line 224C.
23. The hybridoma cell line 224F.

-208-

24. A method of identifying a compound that inhibits M CCS1 comprising the steps of:

- a) expressing M CCS1 in a genetically altered cell, thereby decreasing the sensitivity of the cell to DNA damage, said sensitivity being associated with the genetic alteration;
- b) exposing the genetically altered cell of step (a) to DNA damaging treatment in the presence and absence of a test modulator compound;
- c) measuring the sensitivity of the cell to DNA damage; and
- d) identifying a test compound that restores the sensitivity of the cell to DNA damage as an inhibitor of M CCS1 activity.

25. A method of identifying a compound that inhibits ATM comprising the steps of:

- a) expressing ATM in a genetically altered cell, thereby decreasing the sensitivity of the cell to DNA damage, said sensitivity being associated with the genetic alteration;
- b) exposing the genetically altered cell of step (a) to DNA damaging treatment in the presence and absence of a test modulator compound;
- c) measuring the sensitivity of the cell to DNA damage; and
- d) identifying a test compound that restores the sensitivity of the cell to DNA damage as an inhibitor of ATM activity.

-209-

26. An assay for identifying modulators of ATM kinase activity comprising the steps of:

a) incubating a ATM kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate in the presence and absence of a test compound, and

b) measuring the moles of phosphate transferred to said substrate; wherein an increase in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an activator of said ATM kinase and a decrease in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an inhibitor of said ATM kinase.



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 9/12, 15/63, C07K 16/40, C12N 5/12, C12Q 1/48, C12N 15/11	A3	(11) International Publication Number: WO 97/18323 (43) International Publication Date: 22 May 1997 (22.05.97)
(21) International Application Number: PCT/US96/19337 (22) International Filing Date: 18 November 1996 (18.11.96) (30) Priority Data: 08/558,666 16 November 1995 (16.11.95) US 08/007,312 27 February 1996 (27.02.96) US 08/725,304 21 October 1996 (21.10.96) US (71) Applicant (for all designated States except US): ICOS CORPORATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA 98021 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): HOEKSTRA, Merl, F. [US/US]; 10321 216th Street, S.E., Snohomish, WA 98290 (US). HOLTZMAN, Doug, A. [US/US]; 4308 4th Avenue, N.E., Seattle, WA 98105 (US). KEEGAN, Kathleen, S. [US/US]; 5812 Mercer Way, Mercer Island, WA 98040 (US). (74) Agent: NOLAND, Greta, E.; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606-6402 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> (88) Date of publication of the International search report: 9 October 1997 (09.10.97)
(54) Title: CELL-CYCLE CHECKPOINT PHOSPHATIDYLINOSITOL- (PIK-) RELATED KINASES, GENES CODING THEREFOR AND METHODS FOR ASSAYING AND MODULATING ENZYMATIC ACTIVITY		
(57) Abstract <p>The present invention generally relates to genes encoding cell cycle checkpoint phosphatidylinositol kinase (PIK)-related proteins essential to DNA damage responses in cells. These PIK-related kinases are required in regulatory pathways that arrest the cell cycle following DNA damage to allow DNA repair prior to mitosis or initiation of DNA replication. More particularly, the invention provides a novel human cell cycle checkpoint PIK-related kinase, MCCS1, and polynucleotide sequences encoding the MCCS1. Assays for identifying modulators of MCCS1 useful as, for example, chemotherapy and radiation adjuvants, are also provided by the invention. Further, assays for identifying modulators of the cell cycle checkpoint phosphatidylinositol kinase (PIK)-related protein identified as ATM are provided.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 96/19337

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N9/12 C12N15/63 C07K16/40 C12N5/12 C12Q1/48
C12N15/11

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 09433 A (MEDICAL RESEARCH COUNCIL) 13 March 1997 see SEQ ID No. 1 and 2; pages 16-22; page 33, lines 28-33; page 34, lines 1-12 ---	2-4, 9-14, 16-18, 21,24
P,X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 93, April 1996, pages 2850-2855, XP002023632 CIMPRICH, K.A. ET AL.: "cDNA cloning and gene mapping of a candidate human cell cycle checkpoint protein" see Figure 2 --- -/-	1,3,4, 9-11,15, 16

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- *&* document member of the same patent family

Date of the actual completion of the international search

1 August 1997

Date of mailing of the international search report

26.08.97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Alt, G

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/19337

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>GENE, vol. 119, 1992, pages 83-89, XP002023633 SEATON, B.L. ET AL.: "Isolation and characterization of the Schizosaccharomyces pombe rad3 gene, involved in the DNA damage and DNA synthesis checkpoints" cited in the application see Figure 5; page 84, right-hand column ---</p>	1
P,A	<p>EMBO JOURNAL, vol. 15, no. 23, 1996, pages 6641-6651, XP002023634 BENTLEY, N.J. ET AL.: "The Schizosaccharomyces pombe rad3 checkpoint gene" see the whole document ---</p>	1
A	<p>SCIENCE, vol. 268, 23 June 1995, pages 1749-1753, XP002036686 SAVITSKY, K. ET AL.: "A single ataxia telangiectasia gene with a product similar to PI-3 kinase" see the whole document ---</p>	25,26
T	<p>CURRENT OPINION GENET. DEV., vol. 7, no. 2, 1997, pages 170-175, XP002036687 HOEKSTRA, M.F.: "Responses to DNA damage and regulation of cell cycle checkpoints by the ATM protein kinase family" see the whole document -----</p>	1

Information on patent family members

PCT/US 96/19337

Patent document
cited in search report

Publication date

Patent family member(s)

Publication date

WO 9709433 A

13-03-97

AU 6884696 A

27-03-97